

Navy Personnel Research and Development Center

San Diego, California 92152-7250

TN-97-10 July 1997



Virtual Environment Technology for MOUT Training

Robert F. Helms, II
Daniel B. Nissman
James F. Kennedy
David L. Ryan-Jones

DTIC QUALITY INSPECTED 2

19970811 045

Virtual Environment Technology for MOUT Training

Robert F. Helms, II
Daniel B. Nissman
James F. Kennedy
Research Triangle Institute
Research Triangle Park, North Carolina

David L. Ryan-Jones
Navy Personnel Research and Development Center

Reviewed by
Wallace H. Wulfeck

Approved and released by
J. McLachlan
Director
Classroom and Afloat Training

Approved for public release;
distribution is unlimited.

Navy Personnel Research and Development Center
53335 Ryne Road
San Diego, California 92152-7250

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE July 1997		3. REPORT TYPE AND DATE COVERED	
4. TITLE AND SUBTITLE Virtual Environment Technology for MOUT Training				5. FUNDING NUMBERS Program Element: 0602131M Work Unit: CF31P14.04 Contract No: DAAL03-91-C0034	
6. AUTHOR(S) Robert R. Helms, II, Daniel B. Nissman, James F. Kennedy, David L. Ryan-Jones					
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Navy Personnel Research and Development Center 53335 Ryne Road San Diego, CA 92152-7250				8. PERFORMING ORGANIZATION REPORT NUMBER NPRDC-TN-97-10	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) COMMARCORSSYSCOM 2033 Barnett Ave., Suite 315 Quantico, VA 22134-5010				10. SPONSORING/MONITORING AGENCY REPORT NUMBER	
11. SUPPLEMENTARY NOTES Functional Area: Training Research Product Line: Training Technology Effort: Advanced Training Requirements					
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.				12b. DISTRIBUTION CODE A	
13. ABSTRACT (Maximum 200 words) Military operations in urban terrain (MOUT) is an important component of land combat due to increasing urbanization throughout the world. The objectives of this study were to (1) evaluate the use of VE technology for training military personnel on MOUT tasks, (2) identify the behavioral issues associated with use of VE technology for training, and (3) identify previous attempts to simulate individual behavior in a VE. MOUT were analyzed to determine the applicability of VE technology for training both individual and team-level tasks associated with MOUT. The analysis considered candidate technologies, costs, trade-offs, alternatives, and expected changes in technologies over the next 6 years. The analysis focused on MOUT training tasks that were previously described in U.S. Army and U.S. Marine Corps training manuals, and in previous task analyses performed by other research and development organizations.					
14. SUBJECT TERMS MOUT, virtual reality, simulation, military training, Marine Corps, urban combat				15. NUMBER OF PAGES 163	
				16. PRICE CODE	
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UNLIMITED		

Foreword

This exploratory development effort was conducted under Program Element 0602131M, Project Number CF31P14, Task 4. The work was sponsored by the Marine Corps Systems Command (MARCORSYSCOM), and supports a Marine Corps requirement to design and develop refresher training for Military Operations in Urban Terrain (MOUT). The unique aspects of Marine Corps operations that necessitate this work are that (1) MOUT skills degrade rapidly, (2) Marine Corps ship deployments are long, (3) there are no MOUT facilities for refresher training on-board ships, and (4) there is limited space on-board for the addition of training devices/simulators.

J. McLACHLAN
Director
Classroom and Afloat Training

Summary

Background

As the United States transitions from the Cold War, with its expected battles on the continent of Europe, to a national military strategy based on regional contingencies, the frequency of using military forces is increasing as traditional power structures breakdown and military forces assume new missions such as peacekeeping and famine relief. From the Balkans to Somalia, U.S. military forces are participating in operations that involve the potential for combat. In most of the areas in which U.S. military forces are deployed, they are operating on terrain that has been altered by human activity. Indeed, it is generally agreed that the requirements for conducting military operations on terrain, which has been altered by humans, will continue to increase. Urbanization, the building of cities, towns and villages, represents the most common form of human terrain modification. Thus, Military Operation in Urban Terrain (MOUT) will continue to be an important aspect of land combat training for the foreseeable future.

Objectives

The objectives accomplished during the performance of this work include: (1) evaluating the use of virtual environment (VE) technologies for training military personnel on MOUT tasks, (2) identifying the behavioral issues associated with use of VE technologies for training, and (3) identifying previous attempts to simulate individual behavior in a VE.

Approach

The RTI staff first identified individual and team MOUT training tasks that are considered as candidates for the application of VE. Once these research areas were identified, the RTI research team determined the hardware and software requirements for a testbed to conduct research for individual and team level MOUT training tasks. In the final step for performing this task, the RTI research team determined and recommended the hardware and software requirements for low- and high-cost NPRDC testbeds.

The Research Triangle Institute (RTI) research team then identified the sensory, perceptual, cognitive, motor response, instructional, and training research issues associated with individual and team MOUT training tasks identified in Task 1. The next step in completing Task 4 was to identify candidate VE technologies needed to address those research issues selected for analysis. Once these technologies had been identified, the RTI research team determined the level of maturity and expected changes for the VE technologies over the next 6 years. We then correlated and integrated VE technologies with the requirements for conducting research on the issues identified earlier in this Task. Trade-off analyses were conducted to determine if the applications of VE technologies were needed, or whether there were other more cost-effective methods for satisfying research requirements for these issues.

Last of all, the RTI research team identified research and development projects that have generated algorithms and/or software to simulate a computer-generated environment for motor behavior, sensory and cognitive processes, and/or military tactics of individuals or teams of individuals. As we identified these research and development projects, information and algorithms

for the research and development projects identified were obtained and consolidated. The completion of this work included a traditional literature search as well as an electronic search of appropriate electronic databases and file transfer protocol (ftp) sites.

Conclusions and Recommendations

This research concludes that VEs can be used effectively for training a number of MOUT skills. There remain, however, a number of unanswered questions regarding the optimization of VE training for MOUT applications.

The analysis finds, first, that many MOUT tasks can be successfully modeled and trained in VEs. VE training will be especially useful in the initial states of training, where marines can become familiar with complex, dangerous, and/or expensive equipment and situations, and also in the maintenance or refreshing of skills after they have been learned. For some skills, it appears preferable to conduct most of the training in real-world situations, after a VE "initiation," with follow-up practice in a VE paradigm to maintain skills and to prevent skill decay.

A second finding involves the identification of important areas for future research. Given the newness of VE as a training context, there is uncertainty as to the parameters of effectiveness of simulated environments for training. In particular, there is a need to determine factors that influence the transfer of training from the VE to real-world performance of the task.

An obvious first in VE MOUT training entails creation of a virtual urban environment. Models of cities presently exist that may be adaptable to MOUT requirements. The urban environment should be consistent with USMC MOUT specifications to facilitate the transferring of skills. All buildings should be completely modeled, interior as well as exterior, allowing entry and movement within structures as well as on and between them. Current VE technologies are sufficient for creating the required terrain model of urban areas.

Warfare around anthropogenic obstacles differs from warfare in a forest, jungle, or other natural terrain largely as a function of the presence of people, including friends, foes, those of unknown allegiance, and innocent bystanders. It is, therefore, necessary to populate the urban environment. Artificially intelligent virtual people can be selected from a range of prototypes, from simple preprogrammed repetitive actions to agents with autonomy and intelligence. The environment may also be populated with representations of other participants in a multiuser paradigm.

Many MOUT skills require the presence of solid objects, either to support the trainee or to otherwise apply force against him. It must be acknowledged from the beginning that some tasks can simply not be performed in a VE: for example, one can not climb a virtual rope. However, many of the limitations of virtual reality can be overcome with the ingenious integration of actual objects with virtual ones; for the present discussion, the actual objects are referred to as "workarounds."

The development of a repertoire for workarounds comprises a major effort in itself. While problems can be roughly categorized into types, in a more pragmatic sense each problem will require a specific solution. A few of these are suggested along with a discussion of limitations, which cannot be overcome through this approach. A recommended research program to address workaround issues is also presented more fully in this report.

Contents

	Page
Chapter 1: Introduction	1
1.0 Introduction.....	3
1.1 Background.....	3
1.2 Factors Contributing to Increase MOUT	3
1.3 MOUT Characteristics.....	4
1.4 Nature of MOUT Operations	4
1.5 The Potential for Computer-Generated Simulations.....	5
1.6 Research Purpose.....	6
1.7 Research Objectives.....	6
1.8 Organization.....	6
Chapter 2: Statement of Work	7
2.0 Introduction.....	9
2.1 Task 1.....	9
2.2 Task 2.....	9
2.3 Task 3.....	10
2.4 Task 4.....	10
2.5 Task 5.....	10
2.6 Task 6.....	11
Chapter 3: Major Findings	13
3.0 Introduction.....	15
3.1 Applications of VE for MOUT	15
3.1.1 Movement	16
3.1.2 Entering Buildings	17
3.1.3 Clearing Buildings	17
3.1.4 Establishing Defensive Positions.....	18
3.2 Status of VF Technologies.....	18
3.2.1 Movement.....	19
3.2.2 Sensory Interfaces.....	19
3.2.3 Input Devices	20
3.2.4 Computer Platform.....	21
3.2.5 Software	22
3.3 Areas Requiring Additional Research	22
3.3.1 Software	22
3.3.2 Hardware.....	23
3.3.3 Training.....	24
3.3.4 Integration of VE With Other Techniques.....	24

Chapter 4: MOUT Technology Review.....	25
4.0 Introduction.....	27
4.1 Hardware Systems	27
4.1.1 Visual	28
4.1.1.1 General Information.....	28
4.1.1.2 State-of-the-Technology	33
4.1.1.3 Mapping the Technology to MOUT	35
4.1.1.4 Research Issues	35
4.1.2 Audio.....	36
4.1.2.1 General Information.....	36
4.1.2.2 State-of-the-Technology	37
4.1.2.3 Mapping the Technology to MOUT	37
4.1.2.4 Research Issues	38
4.1.3 Touch	38
4.1.3.1 General Information.....	38
4.1.3.2 State-of-the-Technology	40
4.1.3.3 Mapping the Technology to MOUT	40
4.1.3.4 Workarounds and Research Issues.....	41
4.1.4 Tracking	41
4.1.4.1 General Information.....	41
4.1.4.2 State-of-the-Technology	43
4.1.4.3 Mapping the Technology to MOUT	44
4.1.4.4 Workarounds and Research Issues.....	44
4.1.5 Movement	44
4.1.5.1 General Information.....	44
4.1.5.2 State-of-the-Technology	45
4.1.5.3 Mapping the Technology to MOUT	46
4.1.5.4 Research Issues	46
4.1.6 Input.....	47
4.1.6.1 General Information.....	47
4.1.6.2 State-of-the-Technology	48
4.1.6.3 Mapping the Technology to MOUT	49
4.1.7 Speech Recognition	49
4.1.7.1 General Information.....	49
4.1.7.2 State-of-the-Technology	51
4.1.7.3 Mapping the Technology to MOUT	51
4.1.7.4 Workarounds.....	51
4.1.8 Hardware Platform.....	51
4.1.8.1 General Information.....	51
4.1.8.2 State-of-the-Technology	53
4.1.8.3 Mapping the Technology to MOUT	54
4.1.9 Miscellaneous	54
4.1.9.1 Networks.....	54
4.1.9.2 Physiological Monitoring Systems	54
4.1.9.3 Consequences.....	55

4.1.9.4	Wireless Systems	56
4.2	Software Systems.....	56
4.2.1	VE Tool Kit.....	57
4.2.1.1	General Information.....	57
4.2.1.2	State-of-the-Technology	57
4.2.2	Operating Systems	58
4.2.3	Modeling Packages	59
4.2.4	Autonomous Objects.....	59
4.2.4.1	Simple Behavior.....	60
4.2.4.2	Intelligent Behavior	60
4.3	Recommendations for a Research Testbed	63
4.3.1	The High-End Research Testbed System	63
4.3.2	The Low-End VE Research Testbed System.....	65
 Chapter 5: MOUT Human Factors Issue Review		 67
5.0	Introduction.....	69
5.1	Task 1	69
5.1.1	Fundamental Skills	69
5.1.2	Avoid Open Areas.....	69
5.1.3	Conduct Movement Using Cover	70
5.1.4	Suppress or Obscure Enemy Fires	70
5.1.5	Move at Night or During Periods of Reduced Visibility	70
5.1.6	Select Routes That Will Not Mask Friendly Fires.....	71
5.1.7	Cross Open Areas Such as Streets, Fields, Open Areas Between Buildings, Rapidly Under Concealment of Fires.....	71
5.1.8	Move on Roof Tops That Are Not Covered by Direct Enemy Fires	71
5.1.9	Select Subsequent Positions Before Moving	72
5.1.10	Move Around Corner of Building.....	72
5.1.11	Moving Past First Story Windows.....	72
5.1.12	Moving Past Basement Windows	72
5.1.13	Crossing a Fence or Wall.....	72
5.1.14	Passing or Exiting Doorways.....	73
5.1.15	Movement in Streets	73
5.1.16	Movement Across Open Areas	73
5.1.17	Select, Occupy, and Use a Hasty Firing Position During Movement.....	73
5.1.18	Firing the Individual Weapon During Movement	74
5.2	Task 2.....	74
5.2.1	General.....	74
5.2.2	Fundamental Skills for Entry Techniques.....	74
5.2.3	Entering Upper Floors.....	75
5.2.4	Entering Middle Floors.....	75
5.2.5	Entering Ground Floors	75
5.3	Task 3.....	75
5.3.1	Introduction.....	75
5.3.2	General.....	76

5.3.3	Assign Sectors of Fire.....	76
5.3.4	Eliminate the Threat.....	76
5.3.5	Control the Situation and Personnel	76
5.3.6	Secure and Evacuate Personnel or Equipment While Maintaining Rear Security.....	77
5.3.7	Organization.....	77
5.3.8	Clearing a Room	77
5.3.9	Use C-4, Claymore Mines, and TNT Demolitions to Gain Access to Rooms	77
5.3.10	Clear Rooms with Closed Doors.....	78
5.3.11	Search and Clear Basements.....	78
5.3.12	Avoid Hallways if Possible; If Movement Along Hallways is Necessary, Move Along the Side of Walls as Quickly as Possible to Get OUT of the Hallway	78
5.3.13	Move Between Floors	78
5.3.14	Mark the Building Once it Has Been Completely Secured and Announce "All Secure".....	79
5.3.15	Reorganize the Force	79
5.4	Task 4.....	79
5.4.1	Fundamental Skills	79
5.4.2	Identify and Take Up Hasty Firing Positions	79
5.4.3	Prepare a Defense Fighting Position.....	79
5.4.4	Establish Communications with Other Members of the Defensive Team Via Wire or Radio.....	80
5.4.5	Establish Defensive Fires for Different Weapons for Assigned Sectors	80
5.4.6	Employ Integrated Obstacles, Barriers, and Defensive Fires	80
5.5	Task 5.....	80
5.5.1	Introduction.....	80
5.5.2	Throw Grenades.....	80
5.5.3	Use Camouflage Techniques	81
5.5.4	Booby Traps.....	81
5.5.5	Mines.....	81
5.5.6	Use Obstacles.....	81
5.5.7	Talk With Other Members of the Team Involved in MOUT.....	81
Chapter 6: Research Summary and Conclusions		83
6.0	Introduction.....	85
6.1	Recommended MOUT Training Paradigm.....	85
6.1.1	Population	85
6.1.2	Workarounds.....	86
6.2	MOUT Tasks That are Appropriate for VE Training	88
6.2.1	Task 1.....	88
6.2.2	Task 2.....	90
6.2.3	Task 3.....	91
6.2.4	Task 4.....	91

6.2.5	Task 5	91
6.3	Research Issues	92
6.3.1	Physical Objects in the Actual Environment (Workarounds)	92
6.3.1.1	Force Feedback	92
6.3.1.2	Tactile Feedback	93
6.3.2	Movement of the Participant Through Space	93
6.3.3	Multiple Participants	94
6.3.4	Simulating Human Behavior	94
6.3.7	Explosions	96
6.3.8	Field of View	96
6.3.9	Perceptual Versus Motor Emphasis	96
6.3.10	Summary of MOUT Research Issues	96
References		99
Appendix A--Analysis of Training Requirements for Movement Through Urban Areas OUtside of Buildings		A-0
Appendix B--Analysis of Training Requirements for Entering Buildings		B-0
Appendix C--Analysis of Training Requirements for Clearing Rooms		C-0
Appendix D--Analysis of Training Requirements for Establishing Defensive Positions		D-0
Appendix E--Analysis of Fundamental Skills that are Common to all MOUT Tasks		E-0
Distribution List		

Chapter 1

Introduction

1.0 Introduction

The Research Triangle Institute (RTI) has completed research for the U.S. Navy Personnel Research and Development Center (NPRDC), San Diego to determine applications of virtual environments for Military Operations in Urbanized Terrain (MOUT). The emergence of the suite of technologies that make possible the creation and rendering of realistic, cost effective virtual environments, in combination with the increasing likelihood of urban warfare, make this examination of virtual environment (VE) applications for MOUT very timely and appropriate. The results of this research are presented in this report.

1.1 Background

As the United States transitions from the Cold War, with its expected battles on the continent of Europe, to a national military strategy based on regional contingencies, the frequency of using military forces is increasing as traditional power structures breakdown and military forces assume new missions such as peacekeeping and famine relief. From the Balkans to Somalia, U.S. military forces are participating in operations that involve the potential for combat. In most of the areas in which U.S. military forces are deployed, they are operating on terrain that has been altered by human activity. Indeed, it is generally agreed that the requirements for conducting military operations on terrain, which has been altered by humans, will continue to increase. Urbanization, the building of cities, towns and villages, represents the most common form of human terrain modification.

1.2 Factors Contributing to Increase MOUT

A number of factors are contributing to increased MOUT; and most of these factors are expected to continue into the foreseeable future. Two of the more prominent factors include the population changes that are taking place around the world and the changing nature of low intensity conflict.

The changing populations of the world is increasing the frequency of MOUT. For example, it is projected that, at the turn of this century, nearly three-quarters of the world's population will be living in urban areas located within 500 kilometers of the sea. Not only are the number of people living in urban areas growing and the number of urban areas increasing, the importance of these areas is also increasing with the replacement of agriculture by urban manufacturing and trade as principal sources of national income in third world nations.

These changes have affected the nature of warfare giving rise to the emergence of the urban guerrilla and state-sponsored terrorism as a replacement for the rural guerrilla warfare of the 1950s, 1960s, and 1970s. Guerrilla warfare now takes place in urban areas more often than in the rural countryside. Consequently, the control of urban areas is often the key to control of national resources. Frequently, these urban centers are within striking distance of the sea, or are situated along the sea lines of communication. Early seizure of these areas may support military objectives for follow-on operations.

Additionally, the emerging strategy of the U.S. Navy, which includes increasing focus on military operations across the shore, can be expected to increase the opportunities for marines

to be involved in MOUT. These observations have obvious implications for the U.S. Marine Corps (USMC).

1.3 MOUT Characteristics

While many of the skills required for fighting in traditionally open terrain are transferable to MOUT, a number of distinct characteristics have to be considered when preparing for urban areas. MOUT is characterized by a three-dimensional battle. In addition to fighting the enemy at street level, fighting may also be conducted on roofs and in the upper stories of buildings and below street level in sewer systems, subways, and other underground structures. Assets and resources may be required to deny, retain, secure, or monitor each dimension. It cannot be assumed that the enemy is not there.

The ranges for the employment of weapons and for target acquisition are reduced by urban features. Visibility frequently is limited, and targets may be exposed for only brief periods, often at ranges of less than 100 meters. MOUT involves close, violent combat with considerable reliance on automatic weapons, rocket launchers, hand grenades, and hand-emplaced explosives.

Urban features also increase the difficulty of maintaining effective communications. Tactical radios, the backbone of command and control networks, are limited in range within built-up areas. Consequently, conventional means of communication may be unreliable or only minimally effective.

Conditions for MOUT normally must be expected to favor the defending force. The defender has the advantage of being familiar with the urban area, interior lines of communications, cover and concealment, and the opportunity to prepare features such as ambushes, booby traps, and sniper positions.

MOUT also presents a set of demanding problems for detecting, locating, and identifying enemy forces. The possible presence of civilians who may be indistinguishable from enemy soldiers further complicates the problem of enemy detection. Many methods require line-of-sight, but open areas rarely exist and the enemy is readily concealed in MOUT. In this respect, MOUT typically entails a tremendous advantage for the defender, such that the attacker may require eight or nine times more personnel for operational equality.

1.4 Nature of MOUT Operations

MOUT can involve close combat in multiple directions. Operating from, within, or through urban areas isolates and separates units and marines. The effectiveness of the traditional combat arms team is often reduced. MOUT routinely involves the operations and battles of individual marines, teams, and small units. Marines involved in MOUT must expect to be the target of sniper fire. MOUT is more dependent on the individual marine, team and small-unit leader's initiative, skill, and fortitude than perhaps any other form of military operations.

MOUT regularly involves intense fighting with little or no waning; in possibly no other form of combat are the pressures of battle more intense. Continuous close combat, high casualties, the fleeting nature of targets, and firing from a frequently unseen enemy can be expected to produce

severe psychological strain and physical fatigue, particularly among small-unit leaders and marines.

The requirements for fighting in urban areas make it difficult to transfer skills learned from training for conventional military operations to MOUT. The very nature of MOUT requires training designed to prepare marines to operate under these conditions. The training must include the physical, mental, and psychological demands of MOUT to reduce the stress and to enable marines to contend with the pressures associated with these operations. It must be as realistic as possible, within cost and safety constraints. For example, it should include live fire, troops should wear body armor and use the equipment they will be using in combat, and the training should be physically demanding.

MOUT training should include conditions, such as those described above, that marines can expect to confront in urban warfare. The U.S. Marine Corps has made progress with considerable investment, in developing MOUT training sites. While these sites provide many of the conditions required for realistic training, they are limited in a number of ways. For example, they are:

- costly to build and maintain.
- not transportable, marines must be brought to the location for training.
- difficult to adapt to different settings such as a particular urban area that may become an area of operation.

1.5 The Potential for Computer-Generated Simulations

The emergence of advanced computer-generated models provides opportunities to develop and use cost-effective, realistic simulations to complement existing MOUT training facilities to better prepare marines for the conditions of urban warfare. The most advanced of these models at this time include the ability to generate VEs that approximate conditions in the real world.

Virtual environments result from the integration of a suite of technologies that support the creation of and immersion of individuals into a three-dimensional, synthetic environment; an individual perceives and interacts with this environment and the objects in it in real time as if they were real. VE uses advanced computer hardware and software technologies to create and modify synthetic worlds and objects in these worlds representing different aspects of reality.

That VEs are achievable today is a direct consequence of dramatic technological advances in computer architectures and software systems, and a number of enabling technologies such as data gloves, body suits, voice recognition systems, voice synthesis systems, position trackers, stereoscopic displays, and helmet-mounted displays. Hardware implementations have steadily become more dense (smaller), faster, and less costly. Software systems, too, have become more sophisticated and easier to use. The net result is that powerful workstations with extremely fast CPUs and very sophisticated graphics-rendering capabilities are available at increasingly affordable prices. Moreover, the technological trend is such that, in the foreseeable future, this kind of integrated power will make its way into the high-end personal computers.

Superimposed on this technology base is the industry-wide trend towards standardization of both hardware and software to create increasingly open environments in which workstations of multiple vendors can work cooperatively toward the solution of complex problems. This capability affords scalability of resources, both within a single workstation and across workstations to match the requirements of the applications. Additionally, this capability, coupled with the emergence of user friendly model development systems to support rapid prototyping, modification of models and pictorial databases, and the capability of distributing application software modules across multiple workstations, now makes it possible to use VEs in a number of very practical applications, such as MOUT training.

1.6 Research Purpose

Virtual environments offer the potential for simulations that are transportable and that can be adapted to different settings such as a particular urban area that could become an area of operation. This research was conducted to determine those VE applications that can be used for cost-effective, realistic simulations to support MOUT training for the individual marine and combat team.

1.7 Research Objectives

The objectives accomplished during the performance of this work include:

- evaluating the use of VE technologies for training military personnel on MOUT tasks.
- identifying the behavioral issues associated with use of VE technologies for training.
- identifying previous attempts to simulate individual behavior in a VE.

1.8 Organization

This report contains six chapters and five appendices. This chapter introduces the work completed. Chapter 2 describes the tasks that were completed in accomplishing the research while Chapter 3 presents our major findings. Chapter 4 reviews relevant technology areas, their application to MOUT, and research issues associated with each area in relation to MOUT. Chapter 5 includes our work to identify the sensory, perceptual, cognitive, motor response, instructional, and training research issues associated with MOUT training tasks and to apply relevant technologies to the research issues identified. Chapter 6 presents our conclusions and includes recommendations for using VE in training individuals and teams for MOUT.

The appendices provide a listing of training requirements for the MOUT tasks selected to be included in this work. The tasks selected to be included in this work include: movement outside buildings, entering buildings, clearing rooms, and establishing defensive positions. The final appendix contains a listing of human factor issues that are common to all MOUT tasks.

Chapter 2

Statement of Work

2.0 Introduction

This chapter describes how the RTI staff performed the work included in completing the tasks included in this report. In completing this work, RTI performed five tasks, plus the sixth task of completing this research report.

2.1 Task 1: Analyze MOUT Training Tasks

In completing this task, the RTI research team first identified the training tasks that should be included and analyzed for individual and team MOUT training tasks. The research team conducted a literature search of appropriate materials, including the U.S. Marine Corps (USMC) Combat Review and Evaluation System for MOUT, the U.S. Army's Field Manual (FM) for Combat Operations in Urbanized Areas (FM 90-1), MOUT lesson plans used by the USMC and U.S. Army units, and published articles. The research team also used materials that were available at the NPRDC. The information obtained in this part of the work was used to compile a draft list of individual and team MOUT training tasks.

Once the draft list of individual and team training tasks had been compiled, the Principal Investigator (PI) discussed this list with the NPRDC Contract Officer's Technical Representative (COTR). When there was agreement on the individual and team MOUT tasks to be analyzed, the PI prepared a final list of individual and team training tasks for analysis, and the training requirements that must be fulfilled to perform them were developed and identified.

2.2 Task 2: Determine Applicability of VE Technologies for Training Both Individual and Team Level Tasks Identified in Task 1

In completing this task, the RTI research team identified candidate VE technologies necessary for simulating the individual and team MOUT training requirements identified in Task 1. Once these technologies had been identified, the RTI research team determined the level of maturity and expected changes in the candidate VE technologies as they exist today and are projected to develop over the next 6 years.

The RTI research team developed and adapted a set of criteria to establish categories of priorities for the applications of VE technologies for individual and team MOUT training tasks. Next the research team correlated and integrated VE technologies with the requirements identified for the individual and team MOUT training tasks selected in Task 1. Once the VE technologies had been correlated and integrated with the training tasks to be performed, the research team conducted trade-off analyses to determine if the applications of VE technologies were needed, or whether there were other more cost-effective methods for satisfying these requirements. The RTI research team also identified those areas where additional research is needed to realize the full potential of VE for the individual and team MOUT training tasks selected in Task 1.

2.3 Task 3: Determine Hardware and Software Requirements for a VE Testbed to Be Used for Performing Research on the Issues for Individual and Team MOUT Training Tasks Identified in Task 1

In completing this task, the RTI staff first identified individual and team MOUT training tasks that are considered as candidates for the application of VE. Once these research areas were identified, the RTI research team determined the hardware and software requirements for a testbed to conduct research for individual and team level MOUT training tasks. In the final step for performing this task, the RTI research team determined and recommended the hardware and software requirements for low- and high-cost NPRDC VE testbeds.

2.4 Task 4: Analyze VF Technologies

- Identify the Sensory, Perceptual, Cognitive, Motor Response, Instructional, and Training Research Issues Associated with Individual and Team MOUT Training Tasks Identified in Task 1.
- Determine the Application of VF, Technologies to Research Issues for Individual and Team MOUT Training Issues Identified in Task 1.

In completing this task, the RTI research team began the work by first identifying the sensory, perceptual, cognitive, motor response, instructional, and training research issues associated with individual and team MOUT training tasks identified in Task 1. The next step in completing Task 4 was to identify candidate VE technologies needed to address those research issues selected for analysis. Once these technologies had been identified, the RTI research team determined the level of maturity and expected changes for the VE technologies over the next 6 years. We then correlated and integrated VE technologies with the requirements for conducting research on the issues identified earlier in this Task. Trade-off analyses were conducted to determine if the applications of VE technologies were needed, or whether there were other more cost-effective methods for satisfying research requirements for these issues.

The final step in completing this task was to identify areas where additional research is needed to realize the full potential of VE applications for these issues.

2.5 Task 5: Identify and Describe Research and Development Projects which have Generated Algorithms and/or Software to Simulate the Motor Behavior, Sensory and Cognitive Processes, and/or Military Tactics of Individuals or Teams of Individuals in a Computer-Generated Environment

To complete the work required for this Task, the RTI research team identified research and development projects that have generated algorithms and/or software to simulate a computer-generated environment for motor behavior, sensory and cognitive processes, and/or military tactics of individuals or teams of individuals. As we identified these research and development projects, information and algorithms for the research and development projects identified were obtained and consolidated. The completion of this work included a traditional literature search as well as an electronic search of appropriate electronic databases and file transfer protocol (ftp) sites. The

research team also searched the reports and trip reports of conferences and meetings RTI staff attended and conducted.

2.6 Task 6: Prepare Final Report

The PI compiled the data developed in this work and prepared this research report being provided to the Prime Contractor, the Army Research Office (ARO), and the NPRDC COTR.

Chapter 3

Major Findings

3.0 Introduction

This chapter presents major findings from RTI's study to determine applications of virtual reality that can be used for cost-effective, realistic simulations to support MOUT training for individual marines and combat teams. The chapter discusses feasible applications of VE for MOUT; the status of technology, now and in the future; and the research issues that need to be addressed to realize the full potential of virtual environments for MOUT.

3.1 Applications of VE for MOUT

There are a number of applications where the current VE offers opportunities that the USMC can exploit for significant advantages. Most of these applications are concentrated in the area of acquiring basic principles and skills. That is, VEs can be used for acquiring skills related to sensory, perceptual, cognitive and instructional learning. Conversely, it is difficult, at best, to create VEs that enable the marine to practice motor skills. This examination leads to the conclusion that the USMC should focus the use of VR for MOUT in those areas associated with acquiring basic principles and skills and use other training strategies for practicing skills that are motoric in nature. These strategies may integrate VR into the suite of other training devices, but isolated VEs probably are not the optimal means for practicing skills. Even though the suite of technologies needed to make possible more advanced simulations that could be used for practicing skills is expected to develop over the next 6 years, it is not altogether clear that they will be cost effective. The USMC should keep this issue as an open question and monitor both the development of enabling technologies and their associated costs relative to integrating VEs with less costly, more traditional training strategies.

In the meantime, there are a number of cost-effective MOUT applications that are possible, and steps should be taken to integrate them into MOUT training programs. It is possible to develop virtual worlds that permit an individual to observe and interact with these worlds in ways that enable the individual to become familiar with them, to acquire basic principles and skills from these interactions, and to transfer different levels of learning to real world applications. Examples of these applications include becoming familiar with urban terrain, mission planning, consideration of alternative courses of action, and mission rehearsal.

Virtual environments are particularly well suited for acquiring skills associated with dangerous tasks such as emplacing mines or clearing mine fields. They also can be used effectively for training that otherwise is very costly or that cannot be easily conducted in actual environments. For example, VEs can be used for replicating areas where marines may be employed and used for familiarizing individuals with these environments before they actually begin the operations.

Virtual environments also can be effective as part of the logical progression for traditional learning methods; as an example, the trainee can be provided classroom instruction on clearing a room, then move to experience this task in a VE, and then practice this task at an actual MOUT site. It may even be possible to limit the traditional classroom instruction to an introduction and discussion of principles and use the interactive VE as the principal training method for experiencing the environment before practicing these skills at an actual MOUT site. These models will enable marines to conduct MOUT training more frequently while reducing costs associated

with travelling to distant MOUT sites. They also will reduce the demands being placed on the existing MOUT training sites available to the USMC.

We recommend the use of VE practice in the early stages of training, to introduce marines to unfamiliar skills; further training should subsequently take place in a real-world setting; and finally, refresher training can be conducted in VEs to prevent skill decay.

Equally as important for the USMC, VE models can be transported aboard ship and used for training while the amphibious force is located at sea. It is possible to have database models of different urban areas available that can be loaded in the computer and used for contingency planning, mission rehearsals, and pre-operation training. The availability of these databases would enable marines to plan, practice, and evaluate alternative courses of action. Marines can also visually navigate through the simulation and become familiar with the terrain from different perspectives before the actual operations.

The present report organizes MOUT tasks into categories, which are:

- movement.
- entering buildings.
- clearing buildings.
- establishing defensive positions.

The following section of this chapter describes how VEs can be used for acquiring the skills associated with these MOUT training tasks.

3.1.1 Movement

Virtual environments are particularly effective in helping marines to acquire many of the skills associated with movement during MOUT. While it is difficult to achieve realism in the sense of the marine actually walking through the VE with current technologies, it is possible for the individual to navigate or move through virtual worlds using a device such as a mouse or joystick. These devices are thought to be sufficient for the individual to acquire the basic principles and skills associated with MOUT movement. These tasks, which are particularly movement oriented, can be conducted in a large seasonal room. Also, there are a number of workarounds such as the adaptation of treadmills that can be used to add realism if that is necessary and the additional costs are justified.

Marines can use VEs for navigating through a view MOUT world that requires the individual to avoid open areas, select routes that provide cover and do not mask friendly fires during movement, consider how they would suppress or obscure enemy observation or fires, and move at night or during periods of reduced visibility. Virtual worlds also can be created that enable the marine to cross open areas, move on roof tops, select subsequent positions before moving, occupy hasty firing positions during movement, move around corners, move past basement and first story windows, and cross a fence or a wall. The requirement for marines to pass, enter, or exit doorways also can be modeled in VEs. It is more difficult to achieve realism in a VE for firing the

individual weapon during movement; however, the marine can obtain basic principles for this requirement with present VEs.

3.1.2 Entering Buildings

Virtual environments can be used for obtaining the basic skills and principles for entering buildings. However, it is not possible, at this time, to use VEs to practicing entering techniques such as repelling, use of grappling hooks, and the use of one- or two-person lifts. When weight or force feedback is required, current VEs are not capable of supporting these training tasks. It may be possible, from a technological point, to model these requirements in the future; but it is not at all clear that the use of VEs for such training will be cost effective. It seems there are more cost effective methods for integrating other training devices with VE to provide the desired results. With this caveat in mind, there are a number of training requirements for entering buildings for which VE can be used.

The marine can use VE to conduct a mission reconnaissance of the building to be entered, for mission planning and rehearsal, to acquire the techniques for entering doors and windows, to look for booby traps, and to "see approaches" from the perspective of the enemy defender.

Virtual reality can be used to train individuals in the skills for entering buildings. For example, individuals could search for booby traps, select an entry point, and virtual consequences could be administered. For instance, if the marine fails to search for a booby trap in a likely spot, the computer could generate a warning, perhaps an illuminated "x-ray" view of the concealed booby trap, followed by a simulated explosion, or whatever effect the trap might produce. For more extremely critical work, it may be desirable to inflict pain to simulate the consequences of weapon fire and explosion.

3.1.3 Clearing Buildings

The skills involved in clearing rooms are multifaceted, with motor, cognitive, perceptual, and social components. With appropriate planning, expertise, and technological infrastructure, virtual environments can be used for cost-effective training to acquire the skills required for clearing rooms.

The skills associated with clearing a room include assigning sectors of fire, eliminating the enemy located in the room, controlling the situation and personnel in the room, and securing and evacuating personnel and equipment while maintaining rear security. Virtual models can be used to help individuals search and clear basements, avoid hallways, and move between floors, as well as to mark the building and announce all clear once the building is secure. They also can be used for learning how to prepare and detonate explosives, coordinate movements among team members, control volatile interpersonal situations when rooms are occupied, and move between locations. Virtual simulations can be created to represent a dynamic combat environment, with unexpected events, difficult-to-identify people, randomly placed booby traps, and highly emotional behavior of other actors.

3.1.4 Establishing Defensive Positions

The skills required for establishing defensive positions include cognitive and motor skills. While VEs can be used for cognitive learning, motor skills are best learned in a real world training situation. Consequently, consideration should be given to using VEs for training that is primarily cognitive in nature and using real-world training for motor skills, and then integrating the training at a later time.

Virtual environments are especially well suited for reconnaissance, selection, and planning the construction of defensive positions. The marine in a VE can quickly consider the advantages of various alternatives for defensive positions, to include fields of fire and observation from the perspectives of both the enemy and friendly forces. Different materials for constructing the position also can be considered and evaluated. Obstacles such as mine fields, barriers, and defensive fires can be placed in the virtual world and their effectiveness evaluated.

3.2 Status of VF Technologies

A successful VE system requires that convincing sensory feedback and natural means of interaction be made available to the user. Both sensory feedback and the methods of interaction are effected by a hardware interface. Each sense requires specialized hardware. For example, head mounted displays and image generators are required for vision; and headphones and sound signal processors are required for hearing. Interaction requires its own set of hardware as well: tracking systems to determine the position and orientation of the head and major joints of the body, sensorized gloves for finger position, and speech recognizers for voice interaction.

Two software technologies are also employed in a VE system. The VE tool kit provides an interface between the hardware devices and the model of the VE. The modeler is a separate program used to design all the objects found in the VE.

Some technologies are still in the research and development phase while others are in widespread use. The readiness of the various technologies for inclusion in a VE system depends on the application.

The training goals of the VE MOUT simulator determine how and to what extent the available technologies are used. Conceptual skills and the principles of MOUT can be learned in a simulation employing less sophisticated technology than that required in a simulator seeking to duplicate the actual experience of being in a MOUT situation. Conceptual skills can be obtained while sitting in an armchair. Simulating an actual MOUT experience requires the trainee to move freely. Handling free movement is one of the fundamental bottlenecks in a VE training system. A serious shortcoming in the available technology is in force feedback. It is currently not possible to apply external forces to the body of the kind required to simulate leaning against walls (for example).

The following sections provide a brief overview of the movement issue and the relevant technology areas. The focus of the discussion is on approximating the experience of a real MOUT situation as closely as possible with current technology; conceptual tasks are a subset of the full MOUT experience.

3.2.1 Movement

Allowing trainees to move freely in a VE is a challenging problem. In real life, people move from point to point when they move; in a VE, the individual would soon be outside the volume trackable by the position tracker. There are three ways to solve this problem: using a treadmill, flying through the VE, and using a large work volume.

Treadmills give people the feeling that they are actually walking through virtual terrain without changing position in real space. Unfortunately, treadmills allow movement in only one direction. Despite attempts at making steerable treadmills, they do not provide the free movement required in a VE MOUT simulation.

Flying allows people to move around in a VE relatively freely and in all directions. Direction and velocity are controlled by a hand held input device (such as a Logitech 2D/6D Mouse), by gesture, or by voice command. This method of navigation is appropriate for conceptual training; however, it is not acceptable for a full VE MOUT simulation because the individual's hands need to be free. Moving one's muscles is also more immediate than giving movement commands by voice.

The large work volume approach is the only method that allows free movement. The same problem exists for large work volumes as in the limited movement approaches: one cannot move outside the tracked volume. The larger work volume does allow the trainee to practice certain tasks in a large area. The size of such a work volume could be on the order of 50 by 50 feet and larger. The cost associated with this setup is in setting up the tracking system appropriately. Multiple Ascension Flocks of Birds can be networked together to give the required coverage.

3.2.2 Sensory Interfaces

A convincing VE requires that the senses be presented with high fidelity, artificial stimuli. Currently, only two sensory systems can be presented with reasonably convincing artificial stimuli: vision and hearing. The senses of touch (texture, temperature, proprioception) have received increasing attention with the interest in virtual reality. However, no technology currently exists that is able to reliably present stimuli to these senses. The senses of taste and smell have received little to no attention in the VE community because of a lack of need for virtual taste and smell sensations in current applications; and, a lack of knowledge on how they operate.

Artificial stimuli are presented to the visual system through a head mounted display (HMD). The images are created in an image generator. In a multiple participant VE, each team member must have control over what they are seeing and they must be able to look all around themselves as people do in the real world. Only an HMD approach is suitable for this situation. Current HMDs cannot match the field-of-view (FOV) and resolution of the human visual system. A trade-off exists between FOV and resolution as well. One can expect peak horizontal FOVs of around 100 degrees (horizontal FOV of the human visual system is over 250 degrees) and peak resolutions that approach that of a 1280 x 1024 pixel display on a 17-inch computer monitor on currently available HMDs. Horizontal FOVs greater than 80 degrees seem to be sufficient for the sense of immersion.

The image generator generates images based on the individual's position and orientation in the VE. Image generator performance is typically measured in polygons per second. For the perception of continuous motion, the scene must be updated approximately 15 times per second. Fast, sudden motions require greater update rates. Based on the throughput and the update rate, the maximum number of polygons the image generator can handle can be determined. Current image generators can handle scenes composed of 5,000 (at 30 updates per second) to 10,000 polygons (at 15 updates per second) in the \$100,000 to \$200,000 range. Greater scene complexity is available at greater expense. As a point of reference, a real outdoor scene can contain up to two billion polygons. The application of photographic textures, however, can make the generated scenes look more real.

Three-dimensional (or spatial) sound can be presented to the auditory system through headphones. Making use of the binaural cues for sound localization (interaural intensity and time differences), one can localize sound sources located in front or behind of and to the left or right of oneself. The outer ear alters incoming sounds depending on their elevation. The head-related transfer function (HRTF) models this process. A sound can be placed in the VE, operated on by the HRTF, and then presented to the individual through headphones. The individual can then locate the sound as he or she would in the real world. The technology to implement this approach is mature and is readily available.

Artificial haptic (tactile and proprioceptive) sensations are extremely difficult to present in a convincing manner. There are two reasons for the difficulty in creating artificial tactile (texture and temperature) sensations: lack of a good tactile presentation device and lack of knowledge on how the tactile system works. The proprioceptive (force and body position) system relies on the presence of external forces. The reactive force created by applying energy to the individual's body must be absorbed somewhere. Leaning on a virtual wall is impossible without having something, which must be anchored elsewhere, push back. It is unlikely that force feedback will ever become a reality in VIE simulators without a solution to the anchoring problem.

Finally, taste and smell cannot be fooled in a VE until more research is completed. In particular, the presentation of stimuli to these senses is cumbersome because they are stimulated by chemicals. These chemicals must be stored, combined, and presented to the appropriate receptors with a device that has to cover the individual's mouth and nose. Taste and smell are of questionable use in a MOUT training situation.

3.2.3 Input Devices

There are three types of input devices: position and orientation sensors, hand held devices, and speech recognizers. Position and orientation sensing includes single sensors mounted on a HMD or a glove and joint angle sensors found in sensored gloves. Hand held devices include joysticks and mice. All VE systems using HMDs use position and orientation sensors because they are required to render the appropriate visual scenes. The use of hand held devices and speech recognizers is application dependent.

Tracking systems are used to determine the location of sensors located on one's body. There are several types of tracking systems: mechanical, magnetic, optical, and ultrasonic. Of these systems, the magnetic systems provide the best compromise in characteristics: reasonably large (6

foot cube) working volume, immunity to acoustic noise, and objects intervening between the transmitter and sensor do not affect the sensor output unless the object is metallic. Mechanical systems are very accurate and have very low latencies between position updates; however, one is connected to the system using mechanical arms thus restricting movement. Mechanical trackers are used primarily in tracking HMDs. Optical trackers have the potential for continuous position and orientation estimates; however, they are still in research and development. A class of optical trackers known as video image or pattern recognition trackers use cameras situated around a room and use computer vision techniques to determine the position and orientation of objects of interest. The video image trackers are also still in development. Ultrasonic trackers are popular with start-up VE projects because they are cheap. However, they are susceptible to ambient noise and have higher latency times than magnetic trackers. For the free movement required in a VE MOUT simulator, magnetic trackers should be the choice for present applications.

Sensored gloves and body suits measure the angles of the joints of interest. The most popular device is the sensed glove, which measures the finger joints. In conjunction with a position and orientation sensor, the position and orientation of each finger joint can be determined. This makes gesture recognition by computer possible. Body suits measure the angles of major body joints such as the elbows, shoulders, hips, and knees, thus allowing the whole body to be used as an input device. Sensored gloves are in widespread use. Body suits are only available as custom devices.

Hand held input devices evolved from the joystick and computer mouse. These devices are used as a means of exploring virtual terrain or architecture by flying through the model. The Logitech 2D/6D Mouse and the Space ball are the two most popular such devices. In MOUT, the hands of the trainee must be free. Therefore, hand held devices are not appropriate. For reconnaissance and mission planning, however, VE model fly throughs are appropriate.

Speech is the most natural form of communication for the majority of people. A VE system, which can interpret speech, can make interaction with it more natural. There are a number of commercially available speech recognizers. Some of the most advanced speech recognizers are those based on the SPHINX project at Carnegie Mellon University. SPHINX-based systems are expected to become available commercially in the near future. For MOUT, it is currently sufficient to make use of a large vocabulary, speaker-dependent (system must be trained on the speaker) recognizer.

3.2.4 Computer Platform

The VE simulation is controlled by a computer. This computer, also known as the simulation host, must be able to handle the outputs to all the sensory feedback devices, all the inputs from the various input devices, and control the course of the simulation. Current VE work is based on high-end graphics workstations, such as those made by Silicon Graphics, Inc.

VE applications tend to be made up of several independent processes, which communicate with each other periodically. Examples of processes are the position and orientation tracking module, a spatial sound module, modules controlling autonomous objects, and a glove interface model. Each of these processes affects the load on the computer's processor. The more processes running in the simulation, the more work the main processor has to do. In order to lighten

the load, these processes can be distributed to other computers over a network; or, they can be distributed among processors in a multiprocessor computer. The distribution of processes reduces system latency. The distribution of processes is especially important when designing multiparticipant VEs.

The bottleneck in current VE simulations is in image generation. As a result, the graphics rendering is performed by an image generator (previously described). Simulating intelligent virtual actors and complex physical behavior in a VE requires additional processing power. Multiprocessor computers, such as Silicon Graphics Inc.'s Onyx, can be used to provide this additional computing power.

3.2.5 Software

There are two major software components in a VE system: the VE tool kit and the modeler. The VE tool kit provides the programmer with a set of functions for interfacing to all the input and output hardware described above, maintaining the VE object database, and controlling object interactions. The modeler is used to build all the objects found in the VE.

Simulation software is written with the help of a VE tool kit. The design and purpose of the software, usually written in C or C++, is completely dependent on the programmer; the tool kit simply provides a packaged set of functions for performing tasks needed in most VE systems. Some vendors provide development environments, which reduce the need for writing one's own code. Complex environments, however, will require custom code; development environments can hinder rather than help in these cases. One of the more popular tool kits is Sense8 Corporation's WorldToolKit.

The modeler is not actually a part of the simulation software; it is used to build the VE object database. Modeling is the most time consuming process in the design of a VE. Every object must be described in terms of shape, color, texture, orientation, and physical properties. An object as simple as a tiger might be composed of 2,000 connected triangles. Each triangle has to be characterized. The acquisition of textures to be applied to the faces of the triangles can also be quite time consuming. One of the more popular object file formats is the AutoCAD *.DXF format; most VE tool kits support this format.

3.3 Areas Requiring Additional Research

Virtual reality is a very new technique, and its utility for training applications is being determined. Consequently, there are many research directions to explore. This brief discussion will focus on software, hardware, training effects, and integration of VEs with other training methods.

3.3.1 Software

An urban environment is, above all, a dynamic environment parameters and chains of causation, which might intersect with unpredictable results as the participant interacts with situational objects. We suggest that research should be accelerated toward the development of software implementations of human behavior, as well as toward representing many events that are

common to MOUT tasks. A more thorough discussion of recommended research is presented in Chapter 6.

Human behavior ranges in complexity from simple repetitive actions to fully interactive, agentic interpersonal communications. Virtual actors sampled from any part of this range can have utility; for instance, simple repetitive behaviors can be used to populate a busy marketplace or urban street scene. Therefore, research and development should cover the entire range of potential representations of humans.

This report advocates the use of workarounds, solid objects in the work space, which are integrated into the VE representation. Though specific solutions will be needed for particular problems, such as rifles and walls, a pragmatic research agenda should begin with a taxonomy of problems to be solved, and development of general tools for integration of solid objects in the VE.

Further software issues for VE training include development of realistic explosion algorithms, software to allow multiple participants to interact with one another and with the environment, probabilistic scripting of events, and algorithms to simulate participant movement through space.

3.3.2 Hardware

Virtual reality is essentially a human-computer interface technique; rather than type into a keyboard and look at a television-like monitor, the user inputs data by physically moving, and computer output is usually printed to a graphical display, which the user wears. Thus, the input/output (I/O) hardware should be as sensitive, lightweight, and computationally efficient as possible. Hardware issues that should be addressed in a research program include both I/O aspects of the interface.

Image generators are currently able to render scenes of reasonable complexity (measured in polygons, basic object building blocks). With the application of textures, these scenes can appear more real. However, objects with complex 3D structure (such as animals or machines) cannot be textured to give them realism. These objects need to be composed of a greater number of polygons. Image generators available for under half a million dollars can be expected to render scenes with up to 10,000 polygons at the minimum update rate required for the perception of continuous motion (15 updates per second for each eye). As more realistic scenes are desired, image generators must be able to handle greater scene complexity. Real scenes can easily contain hundreds of thousands of polygons.

Present VEs are lacking in their ability to administer tactile and force feedback to the user. Head-mounted displays are generally dichotomized into those with high resolution and those with wide field-of-view; hardware, which maximizes both parameters, should be developed. This report suggests the integration of workarounds into the VE; these are "hardware" in a literal sense. Methods need to be developed for tracking solid objects so that their representation to the user accurately reflects their position in the real world.

Research should be conducted into the feasibility of inflicting painful consequences for failure to observe appropriate caution in a combat situation. For example, marines who present a

visible profile or expose themselves to enemy fire in a VE may benefit from experiencing painful shocks, as well as other force-feedback consequences.

Lightweight, wireless accoutrements, such as body suits and head-mounted displays, should be developed, as well as a realistic technique to simulate walking across an area.

3.3.3 Training

It is concluded that VEs can be effectively employed for training; indeed, early results are very positive. The parameters that actually affect transfer of training in a VE are not known. Researchers need to investigate the degree of photographic realism, which is required for learning; do objects in the environment need to look real, or can marines learn through interaction with simplified representations? An example illustrating the nature of this question is the human figure model. Realistic human figure models can contain over 20,000 polygons; current image generators can render scenes with up to 10,000 polygons or update rates required for the perception of continuous motion.

A major advantage of VE for training is the ability to participate in seemingly dangerous work, without danger to the trainee. Researchers should question whether VEs actually induce stress in dangerous conditions, as well as how the experience of stress affects transfer of learning. Similarly, VE researchers often use the term "immersion." What is the effect of immersion, or acceptance of the VE as reality, on training? It may be discovered that sophisticated interactive graphics displayed on a CRT result in training effectiveness equal to that of expensive VEs. The research question, which follows from this, is whether it is important to physically perform tasks, as opposed to clicking a mouse and experiencing the perceptual aspect of the task. If users benefit, in terms of training, from armchair participation in interactive graphics, work space and hardware requirements could be greatly reduced.

3.3.4 Integration of VE With Other Techniques

Our perspective is that VE techniques, though they possess many advantages for training, do not hold all the answers; these methods should be integrated with other training techniques to maximize cost-effective learning of skills. Text can be presented in a VE, as can training films and narrative recitation of relevant rules and principles. Natural language processing and voice-recognition methods can be used to facilitate a verbal interface with the system: the user could talk to the computer, ask questions, and receive answers, tips, and important information.

Chapter 4

MOUT Technology Review

4.0 Introduction

A VE is a synthetic world. This synthetic world can be a model of an actual location in the real world or an imaginary place; on a more abstract level, the synthetic world could be a three-dimensional (3D) representation of data from a scientific model. The VE, however, is stored as a collection of numbers in a computer. An interface is needed between the collection of numbers and the senses of anyone wishing to experience the synthetic world. Minimally, one must be able to see, hear, move around, and otherwise interact with the VE. Ideally, one would like to feel, smell, and taste things in the VE. Specialized interface hardware and software is required to make sensory feedback and interaction in a VE possible.

MOUT is an application, which requires sophisticated technology. In MOUT, individuals are required to perform a variety of complex tasks such as locating enemy fire, moving across wide spaces while under fire, entering and securing buildings, and placing and disarming mines and booby traps. These situations require that the members of a MOUT team have a high degree of situational awareness and simultaneously pay attention to detail in their environment. They must also be able to move suddenly, rapidly, and in a variety of postures. These requirements can only be met by sophisticated audiovisual technology and high-speed tracking systems. Current technology is not ready to fully implement a VE MOUT simulator, which could replace actual MOUT training in a real-world environment. However, the technology is at a point where it can effectively be used to acquire certain skills and reinforce skills learned in real world MOUT exercises.

This chapter reviews the relevant technology areas, their application to MOUT, and research issues associated with each area in relation to MOUT. The chapter is divided into four sections: Hardware, Software, Recommendations, and References. The hardware and software sections discuss theory and terminology associated with each area, the current state of the art in the technology, where the technology can be expected to go in the next few years, demands MOUT training will make on the technology, and workarounds that could be used to overcome certain limitations of the technology. The recommendations outline a set of hardware and software components for developing MOUT scenarios and studying implementation and research issues associated with MOUT. Finally, the section on references lists the sources of information used in this chapter--much of the general information is from *Silicon Mirage* by Aukstakalnis and Blatner and from product literature.

4.1 Hardware Systems

The interface hardware provides the physical interface to the VE. The head-mounted display (HMD) and sensed glove are the usual devices associated with current VE systems. There are other system components, however, that are a necessary part of a complete VE system. The hardware can be divided into feedback, input, and miscellaneous components. Feedback components allow the individual to see, hear, and otherwise sense the VE. Input components provide a variety of ways to inform the computer of the person's movements in the VE, as well as ways to interact with it. Each of these system components has a set of issues that need to be considered when putting together a full VE system.

This section will cover visual, audio, touch, tracking, movement, input devices (such as sensed gloves), speech recognition, computer platforms, and miscellaneous hardware. The senses of smell and touch are not discussed because of the lack of technological development in this area. Miscellaneous hardware refers to wireless links, network hardware, physiological monitoring systems, and a means of simulating being shot. For each system, the important characteristics, the current technology, and the mapping of MOUT requirements to that particular system will be discussed. Workarounds and research and development issues will also be discussed when appropriate.

The different hardware systems are at various levels of development. The disparity between the different technology areas will become evident when looking at the availability of these systems in the commercial market. For example, there are a large number of HMDs and image generators available, whereas there are very few tactile and force-feedback devices available. The reasons for these disparities are that there are other industries aside from the VE industry that drive the development of those technologies. As VEs become more attractive from a commercial point of view, those technologies lagging in development are expected to be advanced.

4.1.1 Visual

4.1.1.1 General Information

The visual display subsystem is the most important component of the VE interface. Sighted people have built up years of experience in exploring and navigating the environments in which they find themselves, using sight as their primary sense. As a result, a poor quality visual display system can destroy the feeling of immersion--the feeling of actually being in the VE. Defects in the visual display system will detract from the experience. There are two major components to the visual display system: the display device and the image generator.

There are three display paradigms, or approaches, of interest. The first is flatscreen presentation of stereoscopic images, which are decoded by shutter glasses worn by the user. This provides a good 3D image of the virtual world. However, one must view the screen head-on and the real world is also visible through the shutter glasses. The second approach is based on a projection system. The individual enters a room, which has the virtual world projected on the walls and ceiling of the room. Using shutter glasses, the individual can see 3D objects in all directions. A side-effect of this approach is that objects in the environment with the individual will also appear in the virtual scene and may interfere with the 3D presentation of virtual objects. The view presented can only be based on one participant's orientation. In the third paradigm, the display is mounted on the user's head and the images are presented directly to the viewer's eyes. The individual is able to look all around without restriction of head motion, and objects in the real environment of the user do not interfere with the presentation of the virtual world.

In choosing a display system, it is important to know under what circumstances the system will be used. In the case of MOUT, situational awareness is very important; therefore, head motion cannot be restricted. This requirement reduces the effectiveness of the flatscreen approach. MOUT is a team-oriented endeavor. Each team member must be able to have full control of what is being seen in the environment. As a result, the third paradigm, direct presentation of the images

to the user's eyes, is most viable for MOUT. The HMD is the display technology that implements this approach. Each team member can have his/her own set of images.

The projection-based system of presenting images can sometimes be useful. An instructor or area specialist can give a guided tour of a city and can point out relevant information to the MOUT team. The view of the surroundings is dependent on the location of the instructor in the virtual world. The instructor, therefore, has control over the class or briefing. The CAVE at the University of Illinois at Chicago is a projection-based system (Defanti, Sandim, & Cruz-Neira, 1993).

The image generator (IG) is essentially independent of the display paradigm. The only accommodation required is whether to present stereoscopic pairs for EMD presentation or shutter glass presentation. The IG must be able to handle the scene polygon complexity (the number of polygons in a scene), including their shading and texturing, at the update rates needed.

Head-Mounted Displays (HMDs). Several aspects must be considered when selecting an HMD for an application. Aside from cost, these aspects are horizontal field-of-view (FOV), resolution, depth-of-field (DOF), monochrome versus color, and weight. In general, for a given number of pixels there is a trade-off between field-of-view and resolution. Vertical field-of-view is important as well, but less important than horizontal FOV. A wide horizontal FOV allows HMD wearers to use their peripheral vision. Stimulating the peripheral vision is thought to be an important aspect of the sense of immersion. Depth-of-field describes the ability of the brain to estimate distances.

Field-of-view (FOV). The field-of-view of the human visual system with both eyes facing forward is 180 degrees (250 degrees if the eyes are allowed to move) horizontal by 120 degrees vertical. Of the horizontal FOV, 120 degrees is seen by both eyes. In flight simulator systems, it was found that a 60 to 80 degree FOV was needed for immersion. However, much of the pilot's attention is focussed on what is directly ahead. In simulations such as MOUT where situational awareness is important, a larger FOV is needed to give the participant the same feeling of immersion.

Resolution. Resolution is measured by moving two black lines on a white background towards each other. When there is only a 10 percent difference in intensities between the black lines and the intervening white area, then the resolution of the system has been found. The angular resolution of the human eye in the foveal region (a small region on the retina containing the greatest number of cones, a photoreceptor responsible for the sense of color) is one-half arc minute, or one 120th of a degree. The spatial resolution can be determined given the distance from the eyes. A 17-inch computer monitor viewed at a distance of 46 cm would need a resolution of 4800 x 3840 pixels to match the resolution of the eye (McKenna, & Zeltzer, 1992). This resolution is unattainable with current technology. A more typical resolution is 1280 x 1024 pixels (approximately 1.92 arc minute angular resolution). Some monitors have a resolution of 1600 x 1200 pixels (approximately 1.55 arc minute resolution). Standard VGA resolution on PCs is 640 x 480; at a viewing distance of 46 cm, this is equivalent to an angular resolution of 3.85 arc minutes. Resolution outside the foveal region of the eye falls off rapidly.

Depth-of-field (DOF). There are a number of cues that the human visual system uses to estimate distances. Stereopsis, motion parallax, linear perspective, partially hidden objects, and detail perspective (loss of discernible detail as distance increases) are all important cues (Aukstakalnis, 1992; McKenna, 1992). The cue of greatest interest from a hardware point of view is stereopsis. Stereopsis is the ability of the human visual system to take the slightly different images of an object seen by each individual eye to determine distance. The closer an object is, the greater the disparity in the images. At some distance away, the differences between the objects become negligible. For practical purposes, this distance is around 18 feet from the eyes. In order to simulate stereopsis in the visual interface, each eye must be presented with slightly different images based on the distance of each object in the scene and the image generator must be able to handle double the throughput that a monocular system has to handle. The other cues are monocular cues and are easily handled by proper representation of the physical model of the VE along with an image generator, which knows how to deal with them.

Color. HMDs are available in both color and monochrome. Monochrome display systems are easier and cheaper to build than color displays. Color, however, provides greater realism. Black and white images, also referred to as grayscale images, are used almost exclusively in medical imaging.

Weight. The weight of the HMD should not be ignored. If worn for more than 30 minutes to an hour, the HMD should probably not weigh more than 4 pounds (Latham, 1993a).

There are three basic display system technologies. Each of these systems use some combination of optical elements (lenses, prisms, mirrors) for the final presentation to the eyes. Many of the optical systems in use are based on the Large Expanse Extra Perspective (LEEP) optical design, developed by LEEP Systems, for spreading images from small displays over wide fields-of-view (approximately 120-degrees). In order of least to most expensive, the display system technologies are liquid crystal display (LCD), cathode ray tube (CRT), and fiber optic.

Liquid Crystal Display (LCD). LCD displays are cheap, low-power display devices that are widely used in digital watches and portable computers. They also have the worst resolution and contrast of the three display system types. Resolution is further reduced by the addition of color, which is implemented as three or four color-filtered (red, green, blue, and black) pixels. Manufacturers of LCD HMDs are not consistent in reporting resolution. Most report each of the four color pixels as a separate pixel as opposed to counting "color quads" (Latham, 1993a). Kaiser Electro-Optics is using a novel approach at solving the resolution problem by using multiple LCD displays and using optics to merge them into a high resolution image (Latham, 1993b).

The basic LCD is constructed as follows (Standish, 1991): a thin film of liquid crystal (special rod-shaped molecules that have different electrical and optical properties depending on orientation) is sandwiched between two glass plates. A thin polymer layer on the inside surface of each plate is rubbed in a direction perpendicular to the rubbed polymer at the other surface. The rubbed polymer causes the rod-shaped liquid crystal molecules to align themselves parallel to the direction of rubbing at each surface. A twist in the orientation of the molecules from one plate to the other is the result. Between the polymer layer and the glass plate are transparent electrodes which are shaped to produce the desired image. Polarizers (light filters that transmit light oriented in a particular way) are placed on the outside surfaces of the glass plates and are oriented so the

polarization matches the orientation of the crystals at one surface. If the crystal is not activated, light is polarized according to the first polarizer and is then repolarized through a 90-degree twist by the liquid crystal. Since the polarizer at the other end is perpendicular to the polarization of the light after the twist, light is not transmitted. If the crystal is activated by the application of an electric field, then the crystal aligns itself so that the twist does not occur, and then light is transmitted. The reverse video found in digital watches and portable computer displays is achieved by having the two polarizers perpendicular to each other so that the twisted light is transmitted in the nonactivated state.

There are more advanced LCDs in use, which provide greater contrast and packing densities. Their operation is beyond the scope of this review. However, they ultimately work by manipulating the polarization of light as well. High packing densities are achieved by implanting transistor switching elements directly into the display. LCDs using this technology are known as active matrix LCDs (Standish, 1991; Werner, 1993).

Cathode Ray Tube (CRT). CRTs have higher resolution and higher contrast at the expense of being somewhat bulky and more power hungry than LCD displays. They operate by sweeping an electron beam across a phosphor coated screen. When an electron strikes the phosphor screen, the phosphor emits photons, which are visible to the human eye. The resolution of a CRT system is high because of the very small size of an electron. Contrast is also very good because the energy of the striking electron can be controlled accurately over a wide range. The number of photons emitted from a region of phosphor is proportional to the energy of the impacting electron.

There are several means of obtaining color in a CRT display. One method is to use red, green, and blue phosphors in a grid and three electron guns. As the three guns are swept, the beams they generate pass through a shadow mask, which is positioned very close to the phosphor grid. This mask ensures that the beams hit the correctly colored phosphor. Current HMDs also use monochrome CRTs in conjunction with a color wheel. With this method, for every screen refresh, three sweeps have to be completed—one for each primary color. This corresponds to an effective refresh rate of 180Hz (60Hz is the standard screen refresh rate). The color is set by the color wheel, which rotates the appropriate color in front of the phosphor screen. Since the color persists long enough on the retina, the red, green, and blue components are fused into a single color.

CRTs, found in almost every television and computer monitor, are the most popular display device. Full color HMDs are found mainly in the more expensive models.

Fiber Optic. Fiber optic systems can provide extremely high resolution and contrast in conjunction with a good optical system. Light is projected into each fiber using light-valve projectors (Aukstakalnis, 1992). This light is then imaged at the output through the optical system. These systems are used in conjunction with eye-trackers to provide small, high resolution insets in the displayed regions.

HMDs employing this technology are extremely expensive, ranging from hundreds of thousands to millions of dollars for research devices.

For more information on LCD and CRT displays, see the review paper by Werner (1993).

Image Generators (IGs). Image generators are special hardware devices designed specifically for rendering 3D graphics. They manage displays at the proper perspectives in the VE given the participant's orientation in the VE; they also manage texturing, shading, and scene content. Textures can be applied to object faces to make them look more realistic. In practice, photographs are digitized and applied. Shading is an algorithmic process, which is dependent on the locations of the light sources in the VE. Scene content management allows the generator to maximize the scene complexity without overloading the image generator. The management is done by processing only those elements in the visual database, which can actually be seen by the observer. Image generators developed for simulation systems typically have simulation specific components built-in. For example, the image generators developed for use in SIMNET have added functionality for handling atmospheric effects (haze, fog, smoke), shading and different sensor simulations like infrared scopes.

There are two metrics that are important in evaluating image generators: the polygonal throughput and the update rate.

Throughput. The number of polygons the system can display per second is the system throughput. It is commonly measured in triangles per second. Each manufacturer has a slightly different way of reporting this specification. Some manufacturers measure throughput in terms of textured polygons and some report only flat-shaded (shaded with a single color) polygons. In addition, anti-aliased (see below) polygons are also occasionally included. One must be aware of these differences when comparing IGs from different manufacturers. It has been suggested (Deering, 1993) that all throughput specifications be reported in terms of triangles since all polygons can be decomposed into triangles. Pixels per second is also listed as a peak measure of performance. The pixel content of a scene is dependent on the size of the display and the presence of occluded objects (pixels in occluded objects still have to be processed). Anti-aliasing increases the effective number of pixels as well.

Update Rate. The update rate of the image generation system is the number of fully rendered images it can display per second. Scenes that have large numbers of polygons take longer to render than simple scenes. As a result, the update rate is a function of the scene complexity and the throughput. Motion in complex scenes can appear jerky and unnatural. It is common for the update rate to vary throughout a simulation because of this. In order for the human visual system to perceive continuous motion the system must render an updated scene every 14 to 18 times per second (Deering, 1993). Fast moving objects, however, require higher update rates or else they will appear to jump from point to point across the displayed image. The effects of slow rendering can be lagged images. This lagging can cause the individual to experience cyber sickness, a collection of symptoms that include nausea and fatigue (see Section 4.1.9.2). Image generators have a limit on how high the update rate can be based on hardware considerations.

Update rate is different from the total system latency. System latency represents the total time between a change in state and the rendering of the updated image. The tracking system and image generator are typically responsible for a large portion of this.

High end image generators typically come with a variety of capabilities. These capabilities include z-buffering for handling occluded objects, hardware texture mapping, anti-

aliasing, atmospheric effects, shading, support for multiple light sources, and a variety of special effects. The three most important effects are texturing, anti-aliasing, and shading.

Texturing is the application of bitmapped images onto the surfaces of objects. Texturing systems usually are able to display relatively large (1024 x 1024) bitmaps. Perspective correction is provided by some systems. As a separate option, motion video can be used as a texture. Textures can provide a great amount of realism to a scene; however, they suffer from aliasing effects, which require more processing to correct. Photographic quality texturing is only available on the most expensive systems (\$500K or more).

Aliasing refers to two different things in computer-generated images (Watt, & Watt, 1992). The most common aliasing effect is the jagged line present at high-contrast boundaries. The other occurs when a texture with high-frequency components is subsampled to accommodate the pixel resolution available. In this case, high-frequency information is moved into the lower frequencies, creating an undesirable artifact. The anti-aliasing of jagged lines effectively blurs the edges to make them appear smooth. The anti-aliasing of textures is done through the application of low pass filters.

Shading is the process by which an object is made to look 3D, based on the presence of a light source in the virtual environment. What the viewer sees is dependent on the relative positions of the light source(s), the object, and the viewer. There are a number of shading algorithms of which Gouraud shading is the most commonly used. Some IGs implement Phong shading, a more realistic and computationally expensive shading model.

At the 1st Virtual Reality Annual International Symposium in September, 1993, one mandate of the industry panel to the image generator industry was to concentrate more on increasing the rendering speed for polygons and not worry so much about texturing. In multimillion polygon models, the polygon rendering speed is the bottleneck. For realism in the models, one must go to higher polygon counts (see Section 4.2.3).

4.1.1.2 State-of-the-Technology

Head-Mounted Displays. Almost all commercially available HMDs are based on either LCDs or CRTs. By examining their specifications, the trade-off between FOV and resolution becomes apparent. Almost all high-FOV HMDs are based on LCD technology; whereas, high resolution HMDs are based on CRT technology. The Virtual Research Flight Helmet (\$6,000), based on LCD technology, has a horizontal FOV of 100-degrees and a pixel size of 16.67 arc minutes. The Virtual Reality High Resolution Monochrome Personal Immersive Display (\$47,000), based on CRT technology has a FOV of 60 degrees and a pixel size of 2.82 arc minutes. These two devices represent extremes on the trade-off. There are many devices between these two that can serve as compromises.

There is a wide range of variation, not only in terms of the specifications, but in weight as well. Some HMDs will allow the FOV to be adjusted, which correspondingly affects the resolution. Some HMDs include headphones as an integral part. CRT-based systems are typically an order of magnitude more expensive than LCD-based systems.

Fiber optic systems with eye-slaved insets are expected to be the wave of the future. Currently, the U.S. Air Force Human Resources Laboratory has such a system under development, which will provide a 160-degree horizontal FOV and 1.5 arc minute resolution in the inset. The projected cost of this system will be on the order of one million dollars. The first fiber optic system was built by CAE-Link for the high-end flight simulation industry in Canada. It weighs approximately 5 pounds.

There are some HMDs that superimpose the virtual images onto what is actually visible by the wearer in the real world. A system allowing this is known as an *augmented reality* system and is similar in concept to a head-up display (HUD). The Kaiser Electro-Optics SIM EYE is such an HMD and is touted as being the most cost-effective HMD for military flight simulation at \$95,000. It provides a 60-degree horizontal FOV and a pixel resolution of 2.81 arc minutes. Virtual Reality's High Resolution Color Personal Immersive Display can be fitted for either see-through mode or not.

In the coming years, resolution and FOV will be improved greatly. *Real Time Graphics* predicts that four times the resolution can be expected within the next 2 years. High resolution LCD displays are to be expected in the next few years (Werner, 1993).

For further information on currently available HMDs, see the HMD survey edited by Latham (1993a).

Image Generators. There are two industries driving the image generator market. The military simulation industry requires image generators for flight and tank simulators. The scientific visualization industry requires high-speed rendering of graphical representations of scientific models. The requirements of these two industries are quite different and have resulted in different approaches to the design and packaging of these systems.

In military vehicle simulation, the primary images required are out-the-window scenes. A mock-up of a cockpit or tank is built and computer monitors are placed where all the windows would be. The image generator will then provide appropriate views of each screen. As a result, these image generators handle the interface with the model database directly so that the different views can be rendered appropriately without having to return control to the host computer. In addition, these image generators provide for a lot of atmospheric effects in hardware. For example, Evans and Sutherland's (E&S) ESIG-2000 provides effects for clouds, glare, fog/haze, wet runway, lightning, thunderstorms, and patch fog. Since these simulators are geared to the military, weapon effects are also built in. These simulators also provide simulations for specific sensor types such as infrared and night-vision goggles. These systems are typically outboard systems, which interface to another host. Evans and Sutherland is the leader of this group and has recently been awarded several major contracts. Loral Advanced Distributed Simulation and Martin-Marietta are also in this business.

In specific visualization, real-time interaction with data is required; this means that much of the processing needs to be housed in the host computer. As a result, image generators are typically an integral part of the workstation. The general purpose nature of these systems precludes most of the atmospheric effects and the weapons effects found in the military IGs. To replace those

things, these IGs have many more shading and texturing options. The most cost effective image generators in this group are being made by Silicon Graphics.

Both types have a variety of development environments available; the choice of system one chooses depends on the goals of the project. Military simulation IGs are heavily optimized for the application in which they are used. They are also very expensive; their prices range from \$100K to several million dollars. The Evans and Sutherland ESIG-400 (600K polygons/sec) is over \$500K and the Martin-Marietta SE 2000 (180K polygons/sec) is over \$200K. General purpose IGs are usually much cheaper and have much richer programming interfaces. The Silicon Graphics RealityEngine2 subsystem costs less than \$100K. The RealityEngine2, however, only comes bundled with a Silicon Graphics Inc. (SGI) Onyx, a parallel computer. The low-cost configuration (Onyx/2 with RealityEngine2) is approximately \$160K. The SGI machines, however, have a much more rich set of applications available. They are more suitable for research.

Significant improvements will be made in the next several years in the performance/price ratio. Silicon Graphics claims that they will have an image generator capable of rendering a billion polygons per second by the year 2003. Competition between the military simulation image generator giants and Silicon Graphics will bring costs down.

For further information on currently available image generators, see the image generator survey edited by Latham (1993c).

4.1.1.3 Mapping the Technology to MOUT

In MOUT, situational awareness is important, which means that a high FOV is important. Resolution is reasonably important as well for such tasks as locating mines and booby traps. Depth-of-field is important for judging distances. Stereoscopic displays will be needed. Despite the generally poorer FOV found in CRT-based displays, this is not a limitation of the technology. High-resolution, wide FOV LCD-based systems will be widely available soon, however. The choice of system will ultimately be based on cost. An HMD with an adjustable FOV would help in evaluating which combination would maximize an individual's performance.

The requirements of the image generator are dependent on the scene complexity of the VE model and the update rates needed. Simple models can be rendered very fast on most IGs. It is only when more detail is needed that the polygon throughput becomes more of an issue. In military simulation, 5,000 to 10,000 polygons per frame updated from 15 to 30 times per second is the current norm for high-end image generators. The level of detail needed is a matter for research.

4.1.1.4 Research Issues

Several research issues need to be addressed. The first is how much realism is required to have an effective training simulation. The more realism required, the greater the scene complexity and, as a result, the greater the load on the image generator.

The second issue is whether field-of-view is more important than resolution, given the trade-off between the two in current systems.

Finally, health effects from long-term use of HMDs need to be determined.

4.1.2 Audio

4.1.2.1 General Information

Next to being able to see the virtual world, being able to hear things in it is very important. Objects that make sound can provide information not available through the eyes alone. Objects that are out of sight, whether behind a wall or behind the person, can tell the person of their existence and perhaps may be able to communicate with the individual. Ambient sound can also augment the feeling of immersion and presence in the VE.

The human auditory system is capable of localizing sound in space by using the slightly different sound information received at each ear. The brain uses the interaural intensity difference (IID) as the primary cue for determining location. Depending on the source of the sound, the sound will reach one ear before the other and will be more intense at the first ear. For frequencies below 1500Hz, the brain also makes use of the interaural time difference (ITD). Sound will reach the ear closest to the source first. The cutoff frequency for this effect, 1500Hz, depends on the size of the person's head. Higher frequency sounds result in wavelengths shorter than the person's head and will result in ambiguities. Another method of localization is more subtle. Acoustic shadows are created by objects between the source and the ears. Much like light shadows, acoustic shadows can indicate the source of the sound. Most objects capable of producing shadows affect only the high frequency components of the sound. Low frequency sounds tend not to be affected at all. The ability to localize sound is an aspect of binaural hearing, which describes all the effects of hearing with two ears (Gerber, 1974).

The concepts of interaural intensity and time differences only explain the ability to localize sounds to a person's left or right and forward or behind. In fact, people can localize sound up or down and forward or behind as well. This ability is due to subtle changes in the acoustic signal effected by the pinnae (the cartilaginous portion of the ear lying on the outside of the head). Studies have been completed where miniature microphones were placed on various locations on the pinnae and external auditory meatus (the ear canal). Speakers were arrayed around the listener in all three dimensions. Then each speaker was activated individually to record the response at each of the microphones. It was found that the outputs of the microphones were different depending on which speaker was activated. This lead to the concept of the head-related transfer function (HRTF) (Wenzel, 1992).

There are two methods of presenting subjects with 3D sound. The first method is to array stacks of loudspeakers around the listener. The sounds can be reasonably accurately placed using this method and the HRTF does not have to be taken into account. The drawback is that high quality speakers are expensive and there may be unwanted interference from other sources. The second method is to present the sound directly to the listener's ear through headphones. The advantages are that headphones are less expensive, compact, and there is less chance of external interference. The disadvantage is that the HRTF must be taken into account. The compact nature of headphone presentation makes it a prime candidate for implementation in a virtual reality system.

The HRTF is measured by recording the sounds picked up near the eardrum for a source at varying spots around the head in an anechoic chamber. To present a sound to someone in a VE, the relative positions of the source and head are determined and then the sound from the source is convolved with the HRTT. Then the sound is presented to the individual in the VE through earphones. Psychophysical studies have demonstrated that sounds presented in this way using an HRTF measured in an anechoic chamber give a number of artifacts. The most common of these artifacts is a front-back reversal: a sound whose real position is behind the body is perceived to be in front of the body. Real environments are reverberant. Systems combining the HRTF measured in the anechoic chamber along with a model of the reverberations have a lesser incidence of these artifacts. For absolute accuracy, each participant's HRTF must be measured separately. However, work has shown that using a user-independent HRTF has good performance as well (Wenzel, 1992).

4.1.2.2 State-of-the-Technology

Almost all of the research and development work on presenting 3D sound through headphones has culminated in the products of Crystal River Engineering, Inc. Their current top-of-the-line system is the Convolvotron, which can handle four independent sound sources, reflections from six surfaces, and Doppler shift effects at 16 bits. Multiple Convolvotrons provide extra sources. Their low-end system, the Beachtron, allows two independent sources. All their systems are designed for use with IBM AT compatible personal computers. Interface software is available so that applications on other platforms (such as Silicon Graphics machines) can make use of this technology. Three-dimensional audio systems are also being made available by Focal Point and Advanced Gravis.

Sound sources for these systems can include on-board synthesizers, sample playback from CD-ROM, or speech synthesis. In general, these systems will accept any analog source.

In 6 years, the technology will advance to provide many more sources and reflection models for an equivalent price. As faster computers become available, this area will advance proportionally. The competing 3D audio vendors can be expected to bring prices down as well.

4.1.2.3 Mapping the Technology to MOUT

In MOUT, the team members will need to be fully aware of their situation at all times. They need to be able to hear what is going on in places where they cannot see. As a result, sound is a very important source of information about one's surroundings.

The question of when 2D sound versus 3D sound is appropriate needs to be answered. Heuristically, if there are any potential threats within a few body lengths of the participant at elevations other than the participant's elevation, then 3D sound is appropriate. Otherwise, 2D sound should be acceptable. If the individual can see the sound source, then 2D sound combined with the visual information should give the appropriate localization information. By using a 2D sound source, the computation spent on the HRTF can be saved and perhaps more sources can be modeled.

The number of sources that need to be modeled depends on the simulation. Multiple sources will require more hardware. In cases where there are multiple potential sound sources, but they are only active one at a time, then a single source system can be used. However, having to interact with team members and having random sound events occur will usually preclude this for MOUT use.

4.1.2.4 Research Issues

Two sound presentation issues in relation to MOUT need to be considered. Is transfer-of-training equivalent when working with 2D sound instead of 3D sound? For a typical scenario, what is the maximum number of sources needed? In both cases, the results of the research have the potential to save in hardware costs by reducing the number of 3D sound subsystems required by the simulations.

4.1.3 Touch

4.1.3.1 General Information

Our knowledge of the physical world is enhanced through touch because it gives us information about the shape, size, texture, temperature, weight, and other mechanical properties of the objects around us. When we encounter new objects, we often need to touch them to learn more about them. If we see an object we have experienced in this way, even from only one perspective, then we can recall our memory of the properties of the object. In current VEs, people can see and hear with reasonable fidelity, but they cannot feel the objects in their environment. They must use their own prior experience as a guide when evaluating the physical properties of objects they encounter in the VE.

Interacting with objects in a VE becomes difficult when one cannot touch and feel them. In fact, one cannot know that one has come into contact with an object if the only sensory inputs are sight and sound. Some current VE systems provide sound feedback for when one has touched a button or has touched an object, but this is not the same as actually pushing the button or grasping the object.

An important reference in the real world is gravity. This gives all objects weight, which we can feel when we manipulate them. In a VE, one cannot feel weight or any other force without the senses of touch. Performing actions, which are texture or force intensive, in an environment where there is no touch or force feedback can be quite disconcerting. For a truly immersive VE, the individual experiencing the environment must be able to receive haptic feedback.

The haptic system comprises two related systems: the tactile (texture and temperature) and proprioceptive (force and body position) systems. The tactile system gives us information about the texture, shape, and temperature of objects. The proprioceptive system gives us information about the size, weight, and shape of an object as well as information about the general orientation of our bodies and any forces acting on our bodies. The two systems are interconnected and it is difficult to separate which system gives us what information about an object. The next several paragraphs describe the two systems in more detail.

The tactile system, and to some extent the proprioceptive system, makes use of a set of receptors for obtaining information about how the skin is deformed. These receptors, known as mechanoreceptors, are distributed throughout the various layers of the skin and different ones respond differently to touch stimuli. When a stimulus is presented, the brain integrates the signals from these receptors to determine the location of the stimulus on the body as well as what the stimulus is. Hair provides a lever that amplifies tactile sensation. It is difficult to determine how the receptors work in conjunction with the nervous system to convey the tactile sensations that we feel. Only with more sophisticated experimental techniques will the exact function of the receptors be determined (Cholewiak, & Collins, 1991).

Through subtle changes in the patterns of responses from these receptors, texture can be felt. The size of a small object can be determined by examining the edges and the extent of the object on the hand. Temperature is actually sensed through naked nerve endings whose responses are based on the temperature of the underlying skin.

The proprioceptive system makes use of information gathered from muscles and joints in addition to the receptors that the tactile system uses. When we lean on something, pick an object up, or jump, these receptors tell us where force is being applied to our bodies and how much it is. These are important feedback mechanisms for controlling motor action. How much effort do we need to put into a jump or pick an object up? These questions are answered by the proprioceptive system. To fully escape the real world when in a VE, the proprioceptive system must be fooled.

Creating a virtual haptic display system poses many challenges. It is conceivable that tactile sensation can be simulated with an array of tactors housed in a device that fits directly over the skin. When the virtual tactile sensation is presented, the tactors would be activated. These tactors, however, would have to be small enough to provide the same discrimination that humans can feel. In addition, it would involve wearing something to hold the tactors in place. It may be possible to present texture to the hands only with this approach; it is unlikely this approach would work for the entire body because of calibration difficulties as well as potential discomfort to the wearer.

Fooling the proprioceptive system is even more difficult. The application of forces to the body requires applying energy to the body, which means that the reactive force has to be absorbed somewhere. It is easy to apply force feedback through levers and joysticks because they are anchored at the base. Simulating the sensation of squeezing a ball is more difficult--the anchor would be on the person's body itself. Leaning on a virtual wall is not possible without having something, which must be anchored elsewhere, push back. Simulating the weight of virtual objects has the same problems. Any system that can provide force feedback now is extremely bulky.

The problem of simulating touch sensations is even more complex than is alluded to here (Aukstakalnis, 1992). For example, the sensations of putting one's hands in water, licking one's lips, wringing out a wet shirt, and crinkling up paper all activate a combination of receptors. Stimulating only one or two receptor types is not going to be enough. The bottom line is that not enough is known about the tactile system to understand how to fool it into reporting realistic virtual sensations.

4.1.3.2 State-of-the-Technology

Research in the area of tactile and force feedback is in its early stages. A number of devices have been developed, but few are applicable for MOUT. Those approaches showing promise are mentioned here.

The Teletact System is a texture feedback system based on two gloves. One glove is outfitted with force sensitive resistors and is used to acquire texture data. The other glove is outfitted with miniature air pockets, which can be inflated and deflated rapidly for the presentation of textures acquired with the other glove. This system was developed by Airmuscle, Ltd. in conjunction with the British government.

Instead of using air pockets, the use of electrorheological fluids has been proposed (Monkman, 1992). These fluids change their viscosity with an applied electric field. The change in viscosity can be felt as a tactile sensation.

Making use of the memory metal nitinol, TiNi Alloy has developed a basic tactile stimulation element called a tactor. A piece of straight nitinol wire is heated and then quenched to store this state. The wire can then be distorted in any way. When heated, it returns to the straight wire. Briefly, this feature is used in the tactor by anchoring one end of the wire to a block and then bending it over a raised bar near the end of the block. When the wire is heated by application of an electric current, the wire straightens out. Xtensory and TiNi Alloy have teamed up to make available a package called TacTools--an interface and tactile stimulation exploration package.

The Portable Dexterous Master developed at Rutgers University is a force feedback device designed to work with the VPL DataGlove (see Section 4.1.6). The force feedback is applied to the thumb and first two fingers via computer controlled pistons. The pistons are anchored by ball joints attached to a pad placed in the palm of the hand.

A final force feedback approach worth mentioning is the use of conducting polymers (Lawrence, De Rossi, & Baughman, 1993). These polymers, designed for use in robotic applications, contract with the application of an electric field. Although this approach is still in the early research stages, its application to force feedback merits exploration. Anchors for the polymers could be placed at or around the major joints of the body with bands of polymer connecting them. Then, when someone picks up a virtual object in the VE, its weight could be simulated by appropriately contracting the polymer bands.

None of these technologies are mature enough for commercial development. It is difficult to say when they will be, given the relatively small number of people working in the area of haptic feedback for application to VEs. It is a niche market and not as glamorous as building visual or audio presentation systems.

4.1.3.3 Mapping the Technology to MOUT

In MOUT, individuals are required to carry weapons and other objects, search for and/or set booby traps and mines, fire weapons, hide behind and move obstacles, and generally interact

with their environment in a physical way. Haptic sensation is an important part of the experience for acquiring motoric skills and efforts need to be made to include it in a VE MOUT simulator.

4.1.3.4 Workarounds and Research Issues

Until haptic feedback devices have reached the point of mass production, alternatives need to be considered. The simplest alternative is to use props for objects in the environment with which the individual may have contact. Hand held props have their own weight and texture. Props in the individual's real environment can add realism to the virtual environment but may pose a danger to the individual in the real environment if the prop is not represented in the VE. If movement is restricted to being on a treadmill, then hand held props are an alternative.

The difficulties in implementing haptic feedback raises the question of how the lack of haptic feedback in a MOUT VE simulation will affect learning of the appropriate tasks.

4.1.4 Tracking

4.1.4.1 General Information

Tracking an individual's position in a VE is extremely important. The computer has to know the position and orientation of the head for presenting the correct visual scenes. The positions and orientations of the joints in the individual's body must also be known for correct representation of the participant in the VE. The system must be accurate or else the disparities will result in incorrectly rendered perspectives and, as a result, interactions with the environment are reduced. Incorrect position tracking of joints can hinder human factors work involved in designing work areas such as cockpits.

Position tracking tracks position in x, y, and z coordinates and orientation in roll, pitch, and yaw. As a result, these devices are called six degree of freedom position trackers.

A number of factors must be considered when selecting a tracking device. They are position accuracy, lag time, work volume, occlusion immunity, and connectively to the individual. In current systems, there are trade-offs between these factors that must be considered. The brief descriptions of each of these factors that follow are based on a paper on position tracking by Meyer, Applewhite, and Biocca (1992).

Accuracy. Accuracy measures how close the tracker's readings are to the actual position and orientation of the participant.

Log Time or Latency. There is some finite lag time between updates of the individual's position. If the lag time is long enough, fast movements can make the visual scene lag noticeably and, in some cases, can cause disorientation. For immersion, a lag time of 100 ms or less is desired.

Work Volume. Work volume describes the volume of space in the real environment in which the tracker will work correctly.

Occlusion Immunity. This factor determines whether the tracker can sense through other objects in the environment. If there are other participants in the same work area, occlusion immunity refers to whether tracking is affected by individuals in the way of the one being tracked.

Connectivity to the Participant. Some tracking devices require a mechanical connection to the individual. Some require the placement of sensors on the individual's body. The video image tracker has no connection with the body.

There are several types of tracking devices: mechanical, ultrasonic, magnetic, optical, and video image (Meyer, 1992). Each of these approaches to the tracking problem has its advantages and disadvantages. Some of these types have been implemented in different ways. Only those implementations that are commercially available or have achieved some prominence in the research community will be mentioned here.

Mechanical. Mechanical tracking systems are extremely accurate and have very low lag times. Their primary drawback, however, is that the individual being tracked is restricted to a small region. This method is used primarily for tracking head position and orientation. Mechanical linkages to other parts of the body would be cumbersome and would further restrict movement.

Ultrasonic. Ultrasonic systems work by placing a sound source, which emits ultrasonic clicks, on the object (HMD or glove) to be tracked and then processing the signals from a set of microphones. The differences in the times of arrival at each microphone are used to compute the position of the object. Since this method is based on sound, ambient noise can result in erroneous readings. As a result, this method of tracking is not very accurate and its use is limited to low noise environments. The effect of noise can be mitigated by averaging successive position estimates; latency is increased as a result. Work volume is typically cone-shaped and on the order of 6 feet or so. Occlusion renders measurements unusable.

Another method of tracking using ultrasonics is the phase-coherence method. In this method, a continuous sound is emitted, at different frequencies, from three transmitters on the users RMD. The phases between these signals and their references in the receivers are used to determine position. Phase-coherent systems are only accurate within one wavelength. Movement over larger wavelengths causes ambiguities. Ivan Sutherland, a pioneer in virtual reality technology, used the only phase-coherent system ever built. Its primary advantage is that measurements can be made continuously and, as a result, have a high accuracy and low latency.

Magnetic. Magnetic systems lend themselves to the tracking of a pilot's head in the cockpit of an aircraft. Mechanical systems were too constrictive to the pilot and the cockpits were too noisy for the use of an ultrasonic system (Aukstakainis, 1992). Magnetic tracking avoids these problems. Magnetic trackers operate by sending electric current to three coils of wire (the transmitter) positioned at right angles to each other. This sets up three mutually orthogonal magnetic fields. On the object to be sensed is another set of similar coils; each coil measures the field of one of the transmitting coils. The farther one of the sensing coils is from the corresponding transmitter coil, the weaker the current. The output of the sensing coils is used to determine the position and orientation of the device. The source field can be generated using either alternating current (AC) or direct current (DC). If there are other metal objects in the individual's environment, then changing magnetic fields can create eddy currents in these objects, which in turn affect the

readings by the sensing coils. DC systems avoid this by sampling during the steady-state phase of each measurement cycle. In practical situations, these systems must average multiple readings to get reliable position estimates. As a result, latency is an issue with magnetic systems. Magnetic systems are also immune to occlusion by nonferromagnetic objects. The work volume of a magnetic tracker is on the order of a six-foot cube.

Optical. There are a variety of approaches to optical tracking. The most notable approach is to use a set of cameras mounted on the object to be tracked, which are focused on a grid of light-emitting diodes (LEDs). Based on the images of these LEDs, the position and orientation of the object can be determined. Other systems process the images on objects created by laser light passed through diffraction gratings or by using computer vision techniques on raw image data (described below). Optical systems are extremely fast, but lose accuracy the further the object to be sensed is from the tracking source. Optical systems are affected by occlusion. Work volume is by design--it requires outfitting a room.

Video Image. This sort of system operates by using one or more video cameras to record the objects one is interested in tracking. Using computer vision techniques, the position and orientation of the objects can be determined. The advantages are that the participants do not need to wear any sensors or emitters on them. The disadvantages are that the image processing is computationally extremely expensive and the algorithms for doing the image processing are not of sufficient quality to do a flawless job. Occlusion is a problem, but this can be alleviated by using multiple cameras.

Finally, the choice of system for tracking purposes will depend on the environment in which the simulators will be located. If located in a generally noisy environment, then ultrasonic trackers cannot be used. If there are a lot of magnetic fields or obscuring ferromagnetic objects, then magnetic systems cannot be used.

4.1.4.2 State-of-the-Technology

The most popular tracking systems today are the magnetic trackers. Popularized by Polhemus, this approach has become an industry standard. Polhemus's prime competitor is Ascension Technology. Ascension's trackers use DC field generators as opposed to Polhemus's AC generators. Their Flock of Birds has a work volume of 8 feet and has an update rate of 144 position updates per second for up to 30 sensors. Translational accuracy is 0.1" and angular accuracy is .5 degrees. Ascension provides a means for arraying their transmitters over a large area so objects can be tracked over the entire volume.

Logitech's ultrasonic trackers represent the best of its class as measured by popularity. Work volume considerations as well as interfering noise sources make ultrasonic technology not as useful as magnetic trackers. Logitech's Head Tracker has an update rate of 50 position updates per second and a cone-shaped work volume extending five feet from the transmitter. Translational accuracy is 0.004" and angular accuracy is 0.1 degrees.

In the area of optical tracking, the only systems of note are those being developed at the University of North Carolina at Chapel Hill (UNC-CH), Chapel Hill, North Carolina. These

systems are in the research and development phase. This is an area that should be watched, if not heavily encouraged, because of the real-time potential that optical systems have.

With the availability of high-speed hardware image processors, video image-based tracking will become a possibility. The fact that this tracking method relieves the individual of having to wear sensors on the body should be a driving force in its development. The Princeton Engine, developed at Samoff Research Labs, is a machine that is under development for this and other applications.

One particular tracking system will not be able to handle all particular simulation situations. Improvements in work volume, accuracy, and latency will be realized for all these systems in the next few years. However, real advances in usability will come from developing hybrid systems.

4.1.4.3 Mapping the Technology to MOUT

In performing MOUT-related tasks, individuals will be expected to move suddenly, quickly, and be in a wide variety of body positions. In addition, they will need to use both hands for manipulative tasks. A tracking system must, therefore, have low latency and be able to track multiple sensors on a single person. The system cannot restrict the individual either. The choice of the tracker chosen for MOUT must take into account the basic considerations mentioned here as well as how movement, described in the next section, is going to be simulated.

4.1.4.4 Workarounds and Research Issues

High speed tracking is essential to reducing total system latency. Magnetic trackers are popular because of their immunity to occlusion and their relative speed in comparison to ultrasonic trackers. However, their update rates, although good, need to be improved in order to catch sudden, large excursions of a sensor, such as one mounted on a hand. Hybrid trackers should be investigated. For example, magnetic and ultrasonic phase-coherent systems should be investigated. The slower magnetic system will provide the dead-reckoning of state. The phase-coherent tracker will instantaneously report changes in position. Each time the magnetic tracker reports a new state, the phase-coherent tracker will have a new reference position to work from. Accelerometers and/or gyros can perform the same function that the phase-coherent tracker performs in the hybrid system.

4.1.5 Movement

4.1.5.1 General Information

Walking and running are important aspects of interacting with our real environments. We move through physical space primarily by walking and sometimes by other means of locomotion. The effect of this is that we change physical location. Humans walk forward, turn corners, walk sideways, and run. Simulating these experiences in a VE is quite a challenge.

It is extremely difficult to make navigating in a VE realistic to the individual. One cannot physically move in space without leaving the work volume in which the tracking system

works or reaching the end of the tether to the computing equipment. There are three approaches to this problem. One approach is to allow only motion in one direction. Another approach is to abandon the idea of simulating human locomotion altogether and to fly through the VE. The third approach is to allow the participant to move in a large area.

Motion in one Direction. By using a computer-controlled treadmill, an individual can walk straight through some virtual terrain. Perhaps by using voice commands or some other input device, the individual can change direction. The problem, however, is that the treadmill cannot predict rapid changes in the rate of movement in the individual. This can cause problems for the individual who has "stopped" moving and the treadmill is still running. A final disadvantage of this approach is that movements that are more complex than walking cannot be simulated with this technique. At UNC-CH, bicycle handlebars have been mounted on an unpowered exercise treadmill and the individual turns the handlebars to indicate the desired direction (Brooks et al., 1992). Motion on this treadmill did provide a real feeling of walking, but the exertion required made navigation uncomfortable.

Flying. In this scenario, the individual flies through the VE. Motion is controlled using gesture input, by use of a joy stick, or by voice command. This has advantages for exploring large areas of virtual terrain quickly and one can easily view the VE from different perspectives. Depending on the application, this is either an advantage or disadvantage. When exploring scientific data or terrain, this is a useful method. For training tasks requiring the user to do lots of running or walking, this method is less appropriate.

Large Work Volume. In this scenario, a large room is outfitted with a tracking system able to track everything in its volume. The individual can then move freely within the volume. This includes walking, running, crawling, etc. Multiple people could operate in the same work volume. The disadvantages are that there is still only a finite extent to the size of the physical room, the tethers to the participants would be extremely long, and the tracking system would be extremely expensive. The finite extent of the room was overcome in a similar setup at UNC-CH by allowing a button to be pressed to transport the individual to another virtual area (Brooks, 1992).

It is important to note that in an *augmented reality*, a system where the virtual world is superimposed over the real world, the virtual world must be aligned with the real world. As a result, the individual must be able to move around the real world environment. The treadmill and flying approaches are not appropriate for this situation.

4.1.5.2 State-of-the-Technology

Two approaches that have been used to navigate VEs, while also giving the individual the feeling of walking, are mentioned below. These approaches are the most practical in terms of ease for implementation and realism.

Treadmills, despite their disadvantages, give the individual the feeling of physically moving over some distance. Given the situation where one is only interested in training conceptual skills, navigation can be accomplished using directional gaze and voice commands. Gestures might also be used, except that the individual may be carrying a weapon. There a number of computer-controllable treadmills which are suitable for this task.

Walking in Place. This is a technique used by Mel Slater and Martin Usoh in the Department of Computer Science at the University of London (Slater, & Usoh, 1993). The individual in the VE walks in place. A neural network senses the motion from the tracker outputs to determine when the individual is moving and at what rate. This has a further advantage in that one can emulate other movement actions in place (like walking sideways). Direction of movement can be chosen by changing one's orientation in place.

It is not possible to predict what will be available in 6 years. This area has not received as much attention as other VE technology areas. It is likely, however, that with improvements in tracking system technologies, the large work volume approach will be much more feasible. In addition, there is some speculation that, by using biofeedback techniques, an individual can direct motion through thought.

4.1.5.3 Mapping the Technology to MOUT

MOUT tasks are movement oriented. If it is the purpose of the VE technology to take the place of actual field training, then the problem is going to be extremely difficult. If it is the purpose of the VE training to acquire principles and concepts, then movement is not as significant an issue.

It is not currently possible to handle the complex movements in real space that would be needed in a VE that was to take the place of actual field training. Successful completion of MOUT motoric tasks cannot be achieved using only straight-line motion. Limited to straight-line motion, an individual would not be able to move around corners or avoid obstacles--two essential elements of motoric tasks in a MOUT-type situation. Flying is not a possibility for the simulation of MOUT motoric tasks either because sudden changes in direction are not handled well.

The only approach to the movement problem, which allows the individual the freedom to move normally under MOUT-type circumstances is the large work volume approach. Although expensive, such an approach would allow trainees to work on tasks that do not take large areas. For example, entering rooms, moving past windows, and moving using cover. The primary drawback of having the individual move in real space while experiencing the VE, is that they will be wearing a variety of sensors and other gear which could be dislodged or damaged during the exercises.

Training conceptual skills required for performing well in MOUT-type situations does not require the individual to actually perform these complex movements. In fact, the treadmill and Slater's approach (described in Section 4.1.5.2) can be used here. For exploration of urban, or other, terrain, a Space ball (described in Section 4.1.6) could be used.

4.1.5.4 Research Issues

Making the individuals in the VE feel like they are actually moving through an urban area when they are actually in a corner of a dark room is incredibly difficult to do. In fact, it is currently impossible. The two techniques proposed in 4.1.5.2 need to be studied because they provide potentially inexpensive solutions to this problem. For situations where individuals must feel that they are moving physically, the large work volume approach should be considered.

For a contained system, individuals exploring a virtual reality would ideally never change physical location in the real world while navigating in the virtual world. This can be accomplished by suspending the individual in a harness centered in a frame. With appropriate moving platforms, the individual could walk and run but would never change actual position. These platforms would provide the force feedback of the feet impacting on the ground. The RPI Advanced Technology CyberPod is a first step in this direction.

4.1.6 Input

4.1.6.1 General Information

The visual, audio, and tactile systems are all means by which one receives information about the environment one is in. These are feedback systems. In order to interact with the environment, one needs a means of conveying commands to the computer. The simplest of these input devices is the tracking system (described in Section 4.1.4). Using a sensor on the HMD, the computer knows where you want to look by the position and orientation of the helmet. This section discusses other types of input devices. Sensored gloves and body suits will receive most of the attention here. An assortment of hand held input devices are available. Only the most popular of these will be mentioned.

Humans interact with their environment primarily with their hands. Objects are manipulated and pointed at and communication with other humans is enhanced by or totally enabled by (sign language) gestures. It is natural to attempt to emulate this form of interaction in a VE. By tracking the position and orientation of a hand along with the angles on all the joints associated with the hand, gestures can be input into the computer. The general position and orientation of the hand can be determined using one of the tracking systems (typically magnetic or ultrasonic) previously described. There are several basic approaches for determining the angles of the joints. The two most popular are fiber optics and strain gauges (Aukstakainis, 1992). Other techniques include coating mylar with electrically conductive ink (Mattel PowerGlove) and force sensitive resistors.

Fiber Optics. Fiber optic cables are looped over each knuckle and attached to a control box. Light is transmitted through one end and the intensity is detected at the other end by a photodetector. When a knuckle is bent, light escapes through small cuts in the fiber optic cable thus reducing the intensity of the light detected. Based on the detected intensity, the bend due to the knuckle can be determined. The VPL DataGlove is the pioneer of this technique.

Strain Gauges. Miniature strain gauges are placed over each knuckle and even between fingers to measure the strain due to current hand configuration. Based on the strains detected, the computer can determine what the current hand position/configuration is. Virtual Technologies CyberGlove is the pioneer for this technique.

Body suits are an extension of the idea behind the gloves. By placing sensors on all the major joints, the whole body can be used as an input device.

The other input devices, the Logitech 2D/6D Mouse and Space ball, are not as natural as sensored gloves or body suit. Their development was a natural extension of the mouse used as a

pointing device in almost every modern computing environment. The Logitech 2D/6D mouse can operate as a standard 2D mouse or, with an ultrasonic position tracker, as a 6D mouse. In other words, it will provide position and orientation information. The Space ball, which is offered by a variety of VR vendors, consists of a ball mounted on a base. By moving the ball in the direction one wishes to move, light emitted from an LED array inside the ball makes a particular pattern on the inside of the ball, which is detected by an array of photodetectors. The Space ball's advantage is that it sits on a table and does not have to be moved. Having to pick an object up and keep it in the air, while manipulating the VE, can cause unwanted strain on the arm and hand.

Wearing gloves or holding a mouse can cause strain when used for prolonged periods of time. For certain tasks, such as giving the computer commands, based on gesture or mouse input, voice recognition would be more appropriate. Voice recognition is the subject of Section 4.1.7.

4.1.6.2 State-of-the-Technology

The VPL DataGlove is by far the most popular sensed glove used in virtual reality work. Using the fiber optic system previously described, it reports joint movements for the knuckle and first major joint of each finger including the thumb. Some of these gloves have multiple sets of optic fibers for each knuckle to improve accuracy. Position and orientation of the hand is determined by a magnetic position sensor.

Virtual Technologies' CyberGlove uses the strain gauge method of determining joint angles. Not only does this glove provide angle information for all joints in the hand, it also measures abduction between fingers and angles with respect to the forearm. This glove first gained recognition in a sign language finger-spelling to speech generation system.

The most well-known body suit is the VPL DataSuit. The few that have been made are mostly used in research. The DataSuit is based on the same fiber optic technology used in the VPL DataGlove. Four magnetic position sensors are used to track the head, hands, and the back of the suit. VPL will build DataSuits for custom applications.

Other companies are beginning to market full body sensing systems. These systems can actually be full suits or means of placing sensors over the joints of interest. Virtual Presense, a British company, has recently made TCAS's DATAWEAR line of products available. Their literature claims that up to 96 sensors can be supported. Virtual Technologies has made the technology used in their CyberGlove available in a line of products aimed at measuring other major body joints. The CyberSuit is a full-body sensing system. Virtual Technologies will also make custom sensor wear. Currently, none of these body suits has seen much use, which makes it difficult to evaluate them. Should full body sensing be required, one will need to work closely with one of these manufacturers to ensure that one's needs are met.

In the next 6 years, the sensing technologies are likely to improve and advances will be made in reducing the encumbrance of the suits to the wearer. Interlink, Inc.'s force sensitive resistor (FSR) technology and work underway in the conductive polymer field may result in more accurate and/or cheaper joint angle sensors.

4.1.6.3 Mapping the Technology to MOUT

MOUT tasks are very gesture oriented. Searching for mines, indicating trouble spots to team members, and holding weapons are just a few examples where hand position is important. As a result, the individual's hands need to be sensed. Tasks requiring changes in body position and attitude will require that the major joints in the body be sensed. The 2D/6D Mouse and Space ball are not appropriate or natural for this environment.

Depending on the purpose of the training undertaken in a VE MOUT simulator, one can determine the level of sensor inputs. If the training is to be primarily conceptual, then only sensed gloves are all that are necessary. If physical training is desired, then full body sensing is necessary.

4.1.7 Speech Recognition

4.1.7.1 General Information

Speech is the most natural means of communication available to normal humans and it is therefore natural to design a VE to accept it as an input. Speech will be used to direct aspects of the simulation and, ultimately, to communicate with virtual people. The problem, however, is that speech recognition by computer is not easy. To date, there are no systems that can decipher fully natural speech perfectly for any speaker.

Most successes in speech recognition have come from designing speaker-dependent, isolated word, grammar-constrained systems. The greatest success in speech recognition is the SPHINX project directed by Kai-Fu Lee at Carnegie Mellon, in Pittsburgh, PA. Its speaker-independent, continuous speech accuracy was 71 percent (Lee, 1989). Using contextual grammar, accuracy rose to 94 percent. The concepts of speaker-independence, continuous speech, vocabulary size, and grammar context are described below.

Speaker-independence. Speaker-independence is the extent to which a system can be used, without retraining, by any user. Most systems need to be redesigned or require lengthy tuning procedures to make the system usable by a new user. The SPFHIX system has a very short tuning procedure, which occurs while the system is being used and it is transparent to the user. The accuracies reported above for the SPHTNX system are for truly speaker-independent operation. The error rates are reduced by 5 to 10 percent with speaker adaptation.

Continuous Speech. In natural speech, word boundaries are blurred making it difficult for the recognizer to determine where to start and stop analyzing. High accuracy rates are possible for isolated word speech recognizers where the speaker makes an effort at placing strict boundaries on the words spoken. Placing pauses after every spoken word is very unnatural, but makes it easier for the computer.

Vocabulary Size. The size of the vocabulary that the recognizer can understand is directly related to the recognition accuracy. Large vocabulary size introduces three problems: the number of similar sounding words increases, it becomes more difficult to store information relevant to every word, and database searching becomes more difficult given time constraints. Recognizer

accuracy is affected by each of these problem areas. Similarities adversely affect accuracy because similar sounding words or phrases may be confused for the actual utterance. Large vocabularies make it impossible to hold exemplars for every word in the database because of the memory requirements. As a result, subunits are stored. This creates added complexity in that the subunits must be concatenated appropriately and coarticulation effects are not modeled between subunits. Searching the vocabulary list, or lexicon, is also not a trivial task for large databases. As the lexicon increases in size, it is no longer feasible to do optimal searches in a fixed amount of time. Heuristics must be introduced to direct the searches. The use of heuristics can introduce search errors. A large vocabulary is considered to be 1,000 words. Recognizers with a 20,000-word vocabulary are now becoming commonplace in the research community. The larger vocabulary recognizers rely on grammatical context.

Grammatical constraints. Using information about the previous words recognized makes identification of the current word easier. The lexicon, or dictionary, of recognizable words does not only contain the information required by the recognizer to make an identification based purely on acoustic information, but it also contains information about part of speech and meaning of the word. The constraints allow the analyzer to throw out those words that clearly do not follow based on part of speech. The more complex the grammatical constraints the greater the chance of choosing the correct word. However, introducing grammatical constraints increases processing time.

The complexity of utterances that the computer must understand determines the vocabulary size and the grammar.

Speech is carried by sound. Before a computer can even begin to perform an analysis, the sound must be converted into electronic form by a microphone. Then the electronic representation must be digitized using an analog-to-digital (A/D) converter. At this stage, the raw data must be processed to extract useful features. Common approaches include spectral analysis, linear predictive coding (LPC) analysis, homomorphic (perform a nonlinear transformation on the data--usually a logarithm) analysis, and cochlear modeling. These techniques usually make use of a digital signal processing (DSP) chip. The speech recognition process up to this point is mainly hardware-based. The recognition portion is mostly software-based.

There are a variety of approaches to the recognition problem both in terms of what is to be recognized and the techniques used for recognition. There are systems that attempt to recognize phonemes, phoneme groups, or words. The SPHINX system recognizes groups of three phonemes. The recognition of whole words is only feasible with small vocabulary systems. One or more of the following techniques can be used: dynamic time warping, hidden Markov modeling (HMM), neural networks (NN), and expert systems. Statistically based systems typically have better performance. Grammatical constraints and prosody (pitch, stress, and rhythm) are sometimes used to aid in the recognition process.

A basic system requires a microphone, an A/D converter, and processing software. Feature extraction is usually done in hardware using a DSP chip. As more of the processing can be mapped onto hardware, the recognition process will become faster.

Speech recognition is only one aspect of computer speech understanding. Without a natural language processing system capable of determining the meaning of an utterance, the full value of the speech recognizer cannot be realized. The complexity of utterances that the computer needs to understand determines the sophistication of the natural language processor required.

4.1.7.2 State-of-the-Technology

The SPHINX system represents the state of the art in automatic speech recognition. The concepts developed in this system are now finding their way into commercially available products. CASPAR, a speech recognition package based on the SPHINX system, is available from Apple Computer for their line of computers. Microsoft is working on a commercial implementation of Carnegie Mellon's SPHINX-II system. A multitude of speech recognizers are available on PC and workstation platforms for under \$10K. Many are available for under \$1,000. New systems become available every several months.

In the next 6 years, off-the-shelf SPHINX-based systems should be available. Lots of money has been and continues to be put into research and development in the area of speech recognition. Since this is a very difficult area, it is difficult to predict how recognition accuracy will improve.

4.1.7.3 Mapping the Technology to MOUT

The level of speech understanding required depends on the goals of the MOUT tasks. Giving simple commands to virtual actors does not require a sophisticated system. A speaker-independent, small vocabulary, simple grammar recognizer can be used. One may even be able to use an isolated word recognizer. If interaction with virtual actors gets more complex, however, vocabulary size will increase and grammar constraints will need to relax. Continuous speech recognizers will be necessary.

4.1.7.4 Workarounds

In order to avoid the hardware and computational cost of a speech understanding system, one can have a person interpret for, respond for, and control the virtual actor(s) appropriately. The advantage of this approach is that there would be fewer errors in interpretation and responses would be more appropriate. The disadvantage of the approach is that an additional person(s) must be involved with the simulation system.

4.1.8 Hardware Platform

4.1.8.1 General Information

The hardware platform is the computer, or collection of computers, on which all software is run and to which all devices are connected. In simple terms, a computer is composed of a central processing unit (CPU), memory, and input/output (I/O) circuitry. Each of these components is discussed below. The focus of the discussions is on performance.

Current microprocessor systems are rated in terms of the SPECint92 integer performance benchmark (Hardenbergh, 1994). It is a measure of a system's performance while

running a suite of public-domain applications under UNIX. A SPECint92 of 1.0 corresponds to the performance of a DEC Vax 11/780. There is, additionally, a SPECfp92 for measuring floating point performance. A SPECfp92 of 1.0 corresponds to the performance of the Vax 11/780 with the optional floating point accelerator that was available for it.

CPUS. There are two types of CPUs: complex instruction set computers (CISC) and reduced instruction set computers (RISC). RISC computers have dominated the workstation market because of their higher speeds. RISC designs have a single length for all instructions; most of them execute in a single clock cycle, and there is no indirect addressing--all computations are done in data registers. Performing operations in memory requires communication with outboard memory, which slows operations down, especially since it requires multiple clock cycles. In contrast, CISC designs have variable instruction lengths, many instructions take multiple clock cycles to execute, and they may do some operations in RAM (see below) as opposed to the data registers.

Current generation designs use pipelines and superscalar operation to enhance performance. Pipelining divides an instruction handler into stages so that as one instruction is passed to the next stage, the handler can begin processing the next instruction in the current stage. Pipelining allows the clock frequency to be increased. Superpipelining divides the pipeline into even more stages. Superscalar operation allows multiple instructions to be executed simultaneously on different execution units. Some superscalar designs support only a simultaneous integer and floating point operation, whereas others support multiple integer operations as well. Integer operations make up most of the total operations--85 percent in PC-based applications (Hardenbergh, 1994). As a result, superscalar designs supporting multiple integer execution units will perform better on non-floating point intensive applications.

The data-bus is the physical connection between the CPU and memory. The speed with which the CPU is able to communicate with memory is based on the width of the data path and the clock frequency. It is expensive to increase the bus width, however, because of the extra pins that the packaging needs to support.

Finally, CPU performance is limited by semiconductor manufacturing technology. Two primary factors lead to increased performance: feature size and die size. Feature size refers to the size of the smallest feature that can be etched or implanted reliably in a piece of silicon. Die size refers to the overall size of a piece of silicon. As feature sizes get smaller, a given amount of electric current will change a transistor's binary output state more quickly (Hardenbergh, 1994). This allows the clock rate to be increased. The availability of larger wafers of pure silicon allows for more processing elements to be placed on a single chip. Superscalar designs are easier to implement on larger chip sizes.

At the high end, improved performance can be achieved by using processors in parallel. However, to get peak performance from software in this environment the code needs to be written in such a manner that lends itself to parallelization. A parallel computer's performance on a particular application can be maximized with well-written software.

Memory. All programs and data are stored in memory. There are a number of different types of memory: caches, random-access memory (RAM), read-only memory (ROM), and external

memory. Caches have the greatest effect on CPU performance. Caches pull instructions and data out of main memory and make them available to the CPU while the CPU is executing other instructions. The larger the cache, the greater the likelihood of the cache containing the instruction or data element the CPU requires. RAM is where all data (including program code) that the application requires is stored. If the available RAM is not sufficiently large to hold all the data needed, then the data has to be retrieved from an external memory source (such as a disk drive). ROM is where system start-up code and basic I/O code is stored. ROM is not volatile--it remembers its state even while the power is off. External memory can be any of an assortment of disk drives, tape drives, and CD-ROM drives.

Cache memory has the fastest access times, followed by ROM, RAM, and external memory.

Input/Output. All devices are typically connected to an I/O bus, a physical data path, with which the CPU communicates with other devices such as the video adaptor, disk drives, a local area network, and/or a position tracker interface. In the workstation market, the primary bus is the VME bus. The VME bus's data rate is 50MBytes/sec. Many manufacturers of high-end computers, however, have custom high-speed I/O buses for communication with other products of theirs. Silicon Graphics Inc. includes a high-speed bus with a data rate of 320MBytes/sec.

4.1.8.2 State-of-the-Technology

Several processors are at the forefront of microprocessor performance (Geppert, 1993). The RISC processors are DEC's 220MHz Alpha 21064 (SPECint92 of 130), the MIPS 150MHz R4400SC (SPECint92 of 88), IBM and Motorola's PowerPC 601 (SPECint92 of 85), and Texas Instruments' SuperSPARC (SPECint92 of 80). The CISC processors are Intel's 80486, which has a SPECint92 of 28, and the Pentium, which has a SPECint92 of 68. Of the processors listed here, all the RISC processors have a SPECfp92 greater than their SPECint92, and all the CISC processors have a lower SPECfp92.

Workstation manufacturers making systems being used in VE work are Silicon Graphics, Sun Microsystems, and IBM. Silicon Graphics is the industry leader in high-end graphics workstations. Silicon Graphics Inc.'s flagship system is the Onyx, a parallel computer, which can support up to 24 MIPS R4400 processors. The Onyx also supports the most advanced workstation-based graphics subsystem, the RealityEngine2. The basic two-processor Onyx with the Reality Engine2 graphics system lists for approximately \$160K. The other two manufacturers have targeted their products toward the mid-range market (\$50K and less). Sun's systems are based on the SPARCstation ZX. IBM's systems are based on the RS/6000 (SPEC int92 of 36). Kubota has just entered the graphics workstation market.

The trend in computing is that CPU performance doubles every 2 years. There is dispute about how long this trend can last, however, See Hardenbergh (1994) and Small (1993) for more information on the factors affecting trends in processor performance.

4.1.8.3 Mapping the Technology to MOUT

VE simulations are dominated by graphics computations. As a result, almost all VE systems come with special purpose graphics hardware. CPU performance, although important, does not limit VE applications unless complex system dynamics need to be modeled. Of more interest, however, are multiprocessor platforms. The nature of a battle situation is that there are many different things going on at once. Each one of these can be handled by a separate processor. A system that provides multiple processors is desirable. The inclusion of real-world dynamics will be necessary for realism eventually. A scalable, multiprocessor platform is ideal for a VE simulation system because processors can be added as the complexity of the simulations increase.

4.1.9 Miscellaneous

A number of other hardware considerations need to be taken into account: wireless systems, networks, physiological monitoring systems, and means of applying consequences for incorrectly performed actions. This section is organized differently than the other areas because either there is standard hardware for most of the hardware components mentioned or the components are only in the conceptual stage. Each component is described and how it might be used in a VE simulation system is mentioned.

4.1.9.1 Networks

Network hardware is important in two situations: where computing resources are distributed over a network, and where individuals from different physical sites wish to work together in a VE simulation. The Distributed Interactive Simulation (DIS) is one standard that describes this form of operation. A new IEEE standard (IEEE Standard 1278: Protocol Data Units for Entity Information and Entity Interaction in a Distributed Interactive Simulation) has also been issued recently in this area.

Distributed simulations are heavily used in the military where it is not feasible to bring many people together at one site. The SIMulation NETwork (SIMNET) is a large distributed vehicle-based combat simulator, which has been in operation for a number of years. See Calvin et al. (1993) for more information on SIMNET.

4.1.9.2 Physiological Monitoring Systems

The measurement of physiological parameters of people experiencing a virtual environment is useful for two reasons. First, the measurements provide information on the individual's physiological state. An individual's stress level is a valuable tool in evaluating performance in a training situation. Degrees of stress can be correlated with physiological data; as an individual becomes proficient and more comfortable in performing a task, stress should decrease, which will be indicated via physiological parameters. Secondly, it is useful to monitor general physiological response to being in a VE (Eberhart, & Kizakevich, 1993). Cyber sickness, a collection of symptoms including nausea and eye fatigue which mimics motion sickness, is a poorly understood condition arising from being in a VE (McCauley, & Sharkey, 1992; Biocca, 1992). Physiological data can provide a means of characterizing the condition; it can also provide an early warning system for those more susceptible to cyber sickness.

The physiological measurements might include electrocardiogram (ECG), impedance cardiogram (ICG), blood pressure, heart rate, respiratory rate, and galvanic skin response. These measurements can all be taken noninvasively (the skin is not broken to acquire the data) using readily available equipment. In a research setting, invasive measurements such as blood catecholamine levels can be taken. Much of this equipment is useful only if the individual is tethered directly to the equipment, as is typical in a hospital setting. The additional wiring to the individual may interfere with movement or the operation of other worn devices such as gloves (see Section 4.1.6). Ambulatory monitoring equipment, which reduces the restriction on movement, has recently become available.

Ambulatory monitors are carried around by the individual being monitored. Most are attached to the belt. There are two approaches to ambulatory monitoring. The first is to store all physiological information in the monitor's data storage system. The second is to telemeter the data back to a central monitoring station. The Cardiopulmonary Personal Monitor (CPM) developed at Research Triangle Institute monitors ECG, ICG, and impedance pneumograms. Various measurements are taken on these data streams (such as determination of heart rate) and are stored in memory (Kizakevich, Jochem, & Beadles, 1989). A field study using the CPM in evaluating the effect of volatile organic compounds (VOC) on pulmonary function is given in (Kizakevich, McCartney, Jochem, Raymer, & Pellizzari, 1993). The U.S. Army Biomedical Field Monitoring System, developed at Walter Reed Army Institute of Research, measures ECG, body surface and core temperatures, and activity levels (using accelerometers). This information is then telemetered back to a field command post. In addition to the two monitors mentioned here, ambulatory blood pressure and long-term ECG (primarily used for arrhythmia analysis) monitors are commercially available as well. A monitoring system used in VE applications will likely be a synthesis and expansion of the two approaches.

In MOUT training, the measurement of performance versus stress level, as measured by amount of activity around the individual, is an important tool for evaluating performance and for focusing skills development in certain directions.

4.1.9.3 Consequences

When training to do potentially dangerous tasks, it is necessary to gain some sort of feedback on one's performance in a manner that does not harm the trainee yet does not understate the danger. In combat situations, an incorrect move can result in being seriously injured, perhaps shot. A primary advantage of training people to do dangerous tasks in a simulation environment is that if they make a mistake, they will not be hurt. However, the importance of being careful and doing the task correctly must be conveyed to the trainee or else the usefulness of being trained in a simulator may be called into question.

Two different areas may provide insight on how to approach this question: video games and psychological reconditioning. In video games, being shot typically results in one's death. For those games where the perspective is that of the object being controlled by the player, the screen turns different colors to simulate being blown up or the screen cracks to simulate a crash. There is usually some loud and appropriate noise associated with this. Death is also associated with the end of the game session. In psychological reconditioning, a once common method of altering behavior is shock treatment. Every time the person performs the undesired action, they are shocked.

Eventually, the undesired action is associated with a more undesirable shock and the action is no longer done. These two approaches could be combined to great effect.

A device known as the Self-Injurious Behavior Inhibiting System (SIBIS), developed at the Johns Hopkins University Applied Physics Laboratory, was used to treat autistic children who hit themselves in the head (Newman, 1985). Every time they hit themselves in the head, they would be shocked. After some time, all the children who were on this system were cured of this self-destructive behavior. This system could be modified for use in MOUT training.

In the future, direct stimulation of pain receptors may be possible. When this will be possible, however, cannot be predicted.

Death is the ultimate consequence of an incorrectly performed action. The trainee can easily be informed of their death through visual and audio means. Being shot, however, is not so easy. Shocking the trainee can indicate being nonfatally shot and the amount of shock can indicate the severity of their mistake. By using both audiovisual feedback as well as shocks, the realism of the situation can be enhanced without damaging the soldier.

Being shocked is not a pleasant experience. As a result, it needs to be determined whether this form of consequence adds anything to the learning experience other than pain for the trainee. The area of applying effective consequences in a simulation environment needs to be examined further. In addition, there may be a difference between training for conceptual skills and training for both conceptual and physical skills.

4.1.9.4 Wireless Systems

Wireless systems will allow the direct transmission of information both to and from the host computer and the individual experiencing the VE. The advantage of such a system is clear in the large work volume approach mentioned in Section 4.1.5: the individual will be freed of the cabling connecting the host computer with the hardware being worn. As a result, movement will not be restricted except by tracking constraints.

A wireless system used for a fully interactive VE will need to be able to support a bandwidth of at least 80MHz (1280 x 1024 pixels, stereoscopic, 60 frames/sec) just for the video signal alone. Spatial audio and haptic feedback information will increase the bandwidth requirement. The system will need to be two-way: glove input, speech, and physiological information may need to be sent back to the host computer. An optical or radio frequency (RF) approach is called for.

RPI Advanced Technology is developing a wireless HMD.

4.2 Software Systems

The software system, which controls the participant's interactions with the virtual environment, is typically based around a tool kit. This tool kit is a library of functions that provides interfaces for all the different hardware devices that may be connected to the system, importing model data, and object dynamics. In order to use the tool kit, one must write the application

program in C, which is then run like any other program. The operating system is an integral part of any application. Finally, these tool kits require access to a database of objects that define the VE.

The tool kit and the supporting areas will be discussed below. A final section on autonomous objects is included.

4.2.1 VE Tool Kit

4.2.1.1 General Information

The VE tool kit provides a programmer's interface primarily to a set of routines that control the access to various devices (HMDs, gloves, tracking system) connected to the system and graphical rendering support. Object interactions, such as collision detection, are typically supported, as are routines for importing objects created in modeling packages and for creating simple objects within the environment.

The device drivers provided usually support common devices such as the VPL DataGlove and the Logitech 2D/6D Mouse. An important aspect of a VE tool kit is its ability to be expanded by the user. A new device may be invented by the user or by an outside source. It is desirable to be able to add device drivers to the tool kit without having to get a new version.

The tool kit also provides functions for rendering visual scenes appropriately. These routines inform the renderer of the individual's perspective, what light sources are present, and where, in the virtual world, the participant is located. The rendering can be done in software or in hardware depending on the system.

Collision detection is the most basic form of object interaction. It is also one of the most complex. Each time an object's position changes, the computer must check to see if contact has been made with other objects in the environment. If contact has been made, then some appropriate action is taken. This operation is expensive computationally as the object must check distances to a potentially large number of objects. Simulating moving objects using a model of real world dynamics can be equally complex, if not more so. It is not possible to control large numbers of dynamic objects with current computing resources (Pentland, 1991).

The three-dimensional physical model of the virtual world must be created using a modeling package. The most common object file format in use today is the AutoCAD *.DXF file. Most VE tool kits will support several file formats so the developer can use their favorite modeling program.

4.2.1.2 State-of-the-Technology

A number of tool kits are currently available. The majority are available for IBM PC 80386-based and higher systems. These tool kits include Sense8 Corporation's WorldToolKit (WTK), Division's dVS, Autodesk Inc.'s Cyberspace Developer's Kit (CDK), VREAM, Inc.'s VREAM, and Dimension International's VR Studio. In the UNIX-based workstation environment, Sense8's WTK, Division's dVS, IBM's Virtual Reality - Distributed Environment Construction Kit (VR-DECK), and a variety of packages developed at academic institutions are available. The high

performance required in complex simulations will require the power that only a workstation-based system can provide. Sense8's WTK is the most widely-used of all these packages.

Sense8's WTK is available for both IBM-compatible personal computers and Silicon Graphics workstations. A new version was recently announced for Kubota's new line of graphics workstations. A large number of input/output devices and modelers are supported. On Silicon Graphics workstations, the Performer rendering library is used for faster rendering. An attractive feature of WTK is that it is portable across the platforms it supports. Prototype VEs can be developed on personal computers and then transferred to a workstation. The biggest disadvantage of WTK is that it does not have a means of distributing processes to different machines over a network except through a direct serial line.

Distributed environment concepts are finding their way into VE tool kits. A tool kit built around these concepts can distribute processes to different hosts. For example, two hosts can collect all the input data from two individuals and then transmit the data to the simulation host, which updates the state of the environment. One of VR-DECK's primary design goals is to enable the distribution of processes to other machines. A distributed environment is also supported by Autodesk's CDK (Isdale, 1993).

The advantages of distributed computing are best realized when one has multiple single processor machines with which to work. With a multiprocessor machine, such as SGI's Onyx, the workload is automatically distributed across the available processors. This can significantly reduce the need for a distributed environment. Distributed computing over a network is inherently slower because of communication delays via ethernet or direct serial networks.

In the personal computer market, development environments are available. These are interactive programs that allow the VE developer to create and place objects without having to worry about writing any code. Dimension International's VR Studio is such a package (Hayward, 1993). However, the more complex the environment, the greater the chance that custom code will need to be written.

Division, Inc.'s dVS environment provides the whole gamut of operations from distributed processing, high-level authoring tools, and a programmer's interface. Their high-end systems are usually based on Silicon Graphics workstations. Division's systems are typically sold as fully functional packages and are sold primarily to major corporations. The system is powerful, but is not as popular as Sense8's WTK.

IBM's VR-DECK has an interactive interface for adding and removing objects/modules from the simulation dynamically. The package is in beta testing and will currently only run on an IBM RS/6000.

4.2.2 Operating System

Almost all workstations today run some variant of UNIX. However, UNIX was never designed to be a real-time operating system. Simulating virtual worlds is input/output (I/O) intensive and thus the operating system has to be dealt with each time an I/O operation is performed. I/O slows things down. These problems can be circumvented in some cases, but not in all. At the

1993 IEEE Virtual Reality Annual International Symposium the need for a real-time, possibly VE specific, operating system was emphasized.

4.2.3 Modeling Packages

Modeling is the most time-consuming part of building a VE simulation. Even with high-level authoring tools, each object in the environment has to be described in terms of shape, color, texture, orientation, and physical properties. For a large, realistic environment, this can amount to thousands of objects in the database.

One aspect of the object modeling process that is particularly time consuming is the acquisition of textures. For realism, photographs need to be taken and digitized. The resulting digitized images are then applied to the appropriate objects. Depending on the complexity of the objects, the application of the textures can be time consuming as well. In some cases, it may be sufficient to use algorithmic textures such as fractal-based textures.

A fundamental difference exists between modeling for virtual reality applications and modeling for manufacturing. In manufacturing, detail is extremely important so that the object is accurately rendered in metal or plastic. These models can have hundreds of thousands of polygons in them. The number of polygons can increase dramatically as object detail is increased. For example, in a particular tessellation (the division of an object into a collection of connected triangles or polygons) a low-resolution tiger was composed of almost 2,000 triangles and a high-resolution tiger was composed of almost 15,000 triangles (Deering, 1993). Highly detailed VEs can easily be represented by multimillion polygon models. In fact, a real outdoor scene can have up to two billion untextured polygons in it; intelligent use of textures can significantly reduce the polygon count. In virtual reality, update rate is extremely important. The more complex the objects (measured in polygons) in the environment, the slower the update rate. Thus, there is a trade-off between object detail, or realism, and update rate. Without a simplification algorithm, objects built in a manufacturing CAD system cannot practically be used in a VE. Objects will have to be created from scratch for VE applications. However, Sense8's WorldToolKit provides a simplification algorithm.

There are a number of modeling packages available. The most popular packages are AutoCAD, an object modeler, and 3D Studio, an animation package, from Autodesk, Inc. Multigen, Wavefront, and CATIA (used in the aerospace industry) are examples of other modeling packages. The choice of modeller depends on the needs of the project and preferences of the development team.

4.2.4 Autonomous Objects

This brief section describes the behavior of autonomous objects. At the simplest level, autonomous objects can be traffic lights, which change their lights periodically. At the most complex level, they can be virtual people. Virtual people, also known as virtual actors, can be abstract representations of real people and may exist in the virtual world only to add to the atmosphere. If interaction with a virtual person is required, then the quality of the simulation will need to be higher. These people will be expected to respond appropriately when spoken to or when someone's attention has been focused on them. Simulating human intelligence in virtual actors,

whether in a 3D graphical world or in a text-based virtual world, has been the subject of decades of work in artificial intelligence (AI). Artificial Intelligence is an area that has seen many advances, yet a truly computer-based intelligence has eluded researchers.

4.2.4.1 Simple Behavior

Simple behavior is relatively easy to implement. Objects that perform a sequence of tasks, perhaps with a number of variations depending on the state of the virtual environment, are not difficult to implement. An example of this is a dog that runs around the yard for a certain amount of time, goes and sleeps in its dog house, and then checks its food bowl. If the dog has not been fed, then it may bark and scratch at the door. If it sees its owner, it may run up and lick the owner. A similar scenario can be imagined for a guard at a military installation. For many situations, simple autonomous objects performing sequential tasks are adequate.

4.2.4.2 Intelligent Behavior

The modeling of human behavior requires the solution of two major computational problems: the complexity of the physics for movement of the human form and the poorly-understood complexity of human thinking, including not only decision processes but processing of social information. The human skeletal system is essentially similar to any other kinematic system, with more than 200 degrees of freedom. Joints are constrained to move within limits that are known, but the movement of these joints is effected by a nonlinear dynamic system of continuous states of tension and relaxation by the muscles. This flesh is in turn covered with soft tissue, subcutaneous fat and skin, which deforms with movements: all of this results in a very complex system for realistic modeling. Thus, the computation required for realistic representation of human movement is rather staggering.

Realistic representations of human figures in VE applications pose problems for graphical rendering systems. Current image generators can handle scenes composed of on the order of 10,000 polygons at the minimum update rate of 15 updates per second required for continuous motion. Low resolution representations of human figures can contain on the order of 2,000 polygons whereas high resolution representations may contain 20,000 or more polygons in them. For the rest of the scene to appear realistic, the VE cannot currently be dominated with human figure models. Image generator technology must be advanced significantly before realistic human figures can populate VEs. If cost is no object, however, one can spend well over a million dollars for a single image generation system, which will handle perhaps up to four times the scene complexity of image generators available for less than half a million dollars.

Representation of the basic physics of the human form in a realistic environment is not fundamentally different from the representation of inanimate objects. According to Badler, Phillips, and Weber (1993), the forces and moments which are most important to consider in modeling human figures include:

- gravity, acting upon each segment of the body.
- internal forces generated by muscles: modeled as a driving moment at the joint.

- reaction forces generated by the body's surroundings, for example when the body leans on something.
- external forces, for instance objects that are moved by the figure.
- collisions.

The problem-solving required to perform basic human tasks is staggering, as well. For instance, Zeltzer and Johnson (1993, in press) are working on programming a virtual actor to "Go to the kitchen and get me a beer." Their description of the hierarchies of motor tasks involved in such a simple behavior suggests the difficulty of programming an agent who is capable of forming complicated decisions and executing behaviors based on those decisions in real time.

The problem-solving aspects of a virtual actor resemble aspects of the problem-solving requirements of robotics. The difference, of course, is that the robot is constrained by physics, while physical laws need to be written into the VE program, to the extent that the simulation is intended to mimic reality.

The interactive behavior of virtual actors can be made more complex and more human-like by incorporating natural language processing, expert systems, knowledge representation, and problem-solving techniques into their programming. As the level of behavior expected from the autonomous object increases, the amount of processing by the computer increases as well. The database searches alone can be quite expensive. Reasoning based on the results of the searches is more complex as it may require multiple queries as it gets closer to a solution. In short, complex behavior is attainable at the expense of computing resources.

In the computationally intensive environment of a VE simulation system, one must be careful to balance functionality with the goals of the simulation.

Major research programs in simulation of human behavior are being conducted at the Massachusetts Institute of Technology (MIT) and University of Pennsylvania (UP), with some collaboration among researchers at various other institutions. The two major programs will be overviewed here, with mention of recent breakthroughs and features that distinguish the two approaches. According to (Latham, 1993d), however human figure modeling is "in its infancy." More work will be required to create convincing, human interactive figures.

Massachusetts Institute of Technology. The MIT research team, led by David Zeltzer, is developing a virtual actor they have named "Dexter." A software collection called "WavesWorld" has been created for designing, building, and debugging simulations containing, among other things, virtual actors. WavesWorld is integrated with the NeXTSTEP development environment, which acts as a direct-manipulation, or guiding front-end to multiple, heterogeneous computing resources.

In WavesWorld, an actor consists of three kinds of objects (instances of superclass ActiveObject), called BodyManager, Planner, and AgentManagers. BodyManager is an active object that maintains the internal (energy level, likes and dislikes, etc.) and external (geometric, kinematic, and dynamic properties) states of the virtual person. Planner executes the planning

algorithm, calling skills and updating local information concerning the various mechanisms that implement and control the virtual actor's behavior. Finally, AgentManagers manage goals, skills, and perceptions: a single virtual actor may have many AgentManagers.

The MIT research effort focuses on control of the virtual actor at the task level. Task-level interaction allows the virtual actor to respond to spoken language (this laboratory seeks to develop a natural-language interface for virtual actors in virtual environments) or text input. As has been widely reported, however, the specification of human tasks in quantitative or qualitative terms is very difficult. Thus, an interface agent called TaskManager translates text or verbal task-descriptive input into task primitives, which are well-defined motor units that can be named and simulated, such as grasping, walking, etc.

The MIT lab has also taken on the interesting task of simulating human facial expression (cf. Chen, Dickens, Gaines, Metzger, Miller, & Owen (1993). Though various methods have been used to animate facial representations, the most recent work published by these researchers implements a technique based on Pieper's finite element method (FEM). According to them, the advantages of this method of modeling soft skin and soft tissue of the face include the accuracy of the model, which is based on a video scan, and the efficiency of calculations.

The MIT scientists model skeletal movement mainly as a kinematic rather than dynamic system. Jointed figures are considered to be networks of linked manipulators, with limbs attached to a reference point (the body), and hands and feet cast as end effectors. The disadvantage of kinematic modeling is that it cannot account for the dynamics of the systems involved. The MIT modelers have, however, included dynamic considerations in the programming of the eyes of the virtual actor. Voluntary saccades and smooth pursuit of objects are programmed to respond to objects in the virtual environment according to experimental data from measured responses of human subjects.

University of Pennsylvania. Researcher Norman Badler has been investigating techniques for representing human movement in computer graphics since the mid-seventies. He and his colleagues have developed a software system called "Jack." Jack incorporates algorithms for anthropometric human figure generation, a flexible torso, multiple limb positioning under constraints, view assessment, reach space generation, and strength-guided performance simulation of human figures within a three-dimensional VE. An important consideration for the designers of Jack was the creation of a figure that moved realistically, similar to an actual human body. The creators of Jack also allowed for high-level task control with various knowledge bases, task definitions, and natural language instructions, built into an interface that is intuitively simple to use, yet very versatile.

The representation of Jack is intended to establish a compromise between depictional realism and display speed. The body is composed of 69 segments with 68 joints, and is made (including cap and glasses) of 1,183 polygons. The same researchers have experimented with more realistic-appearing bodies, scanning 89 subjects and transforming the images to polygons. As real humans are not symmetrical, only data from the right halves were used; the resulting figures had 39 segments modeled with about 18,700 polygons.

Jack is operated by manipulation of a database called Peabody, which represents figures composed of segments connected by joints. Peabody contains information about segment dimensions and joint angles, and also efficiently computes and manages geometric information.

The researchers at the University of Pennsylvania have not concerned themselves with modeling of facial expressions, soft tissue, or deformable clothing. For instance, when a joint is bent, a "gap-filling" algorithm is employed to fill the gaps between rigid segments. They have rather emphasized the importance of modeling a realistic spine, so that the torso can bend in accurate imitation of the human body. A database maintains the unique set of features for an individual's spine and torso, including degree of flexibility, size of vertebrae, joint limits, joint rest position, and range of movement. Medical data provide the parameters for the average person.

4.3 Recommendations for a Research Testbed

The components comprising the research test bed system recommended in this section were chosen based on the following principles: the VE MOUT simulator will accommodate multiple participants; research questions based on how the technology can best be used in MOUT can be addressed appropriately, and expendability. MOUT is a team-oriented endeavor; multiple participants must be supported in a VE MOUT simulator. The answers to the research questions posed in this chapter will provide the specifications of VE MOUT simulators that could ultimately be placed in the field. It is important that the research system be capable of addressing the research questions. If it is discovered that the system chosen does not have the performance required to run a VE MOUT scenario, then, in general, it is cheaper to expand a current system than to buy a new one.

This section is divided into two parts: a high-end research system and a low-end research system. The high-end system will allow full-realization of the questions posed in this chapter. The low-end system will be a scaled back version of the high-end system.

As new and improved technologies become available, the specific equipment recommendations made here will change.

4.3.1 The High-End Research Testbed System

The high-end research system will be based on four participants. Each participant will require HMD, headphones, gloves, and tracking. Of greatest importance, however, is that each participant will require an image generator. The image generation system must be able to render a given scene at an update rate that will cause motion to appear continuous. An update rate that appears to be acceptable (Brooks, 1992) is 15 updates per second for architectural walk throughs. The conservative 300K polygons/second rendering rate for the SGI RealityEngine2 translates to 10,000 polygons maximum scene polygon complexity at an update rate of 15 updates per second per eye for stereoscopic viewing. This will provide reasonable detail with texturing, but it is minimal. More complex scenes will slow the rendering rate.

The RealityEngine2 graphics subsystem is only available with the Silicon Graphics Onyx series of parallel graphics supercomputers. Fortunately, the Onyx is a scalable computer--it will be available in 2, 4, 8, 16, or 24 processor versions. As the complexity of ones simulations

increase, more processors can be added. In addition, the graphics system can be scaled as well. More RealityEngine2 subsystems can be installed and RasterManagers, systems which improve pixel rendering performance, can be installed as well. The combination of the SGI Onyx and RealityEngine2 has become the industry standard high-end virtual reality workstation.

It is recommended that three Silicon Graphics Onyx/2's and one Silicon Graphics Onyx/4 serve as the computer platform. Each Onyx will handle the sensory feedback and input from each participant. The Onyx/4 will serve as the simulation host. Each Onyx will house a RealityEngine2 graphics subsystem, a Multichannel board (required to provide stereo images), and two RasterManagers. The cost for the Onyx/2s is approximately \$240,000 each. The cost for the Onyx/4 is approximately \$300,000. Should it be decided that graphics performance is lacking, more RasterManagers can be added or even another RealityEngine2 subsystem. More processing power is available with the addition of extra processors. The inclusion of virtual actors in the simulations will require additional processing power because of their computationally expensive nature.

Requirements of the HMD to be used in the research system are reasonably high resolution with a maximum FOV over 100 degrees, a variable horizontal FOV, color, and a weight under four pounds. The RPI Advanced Technology Head Mounted Sensory Interface (~\$50,000) has a resolution of 3.84 arc minutes (equivalent to VGA resolution of 640 x 480 pixels) at an 82 degree horizontal FOV. The FOV is variable up to 110 degrees. It is color and weighs only 4.5 ounces. An added bonus is that the RPI HMD is available with a wireless link, high fidelity headphones and microphone system.

The spatial (3D) audio system that is recommended is the Crystal River Engineering Convolvotron (~\$15,000). It represents the state-of-the-art in headphone presentation of 3D sound. The Convolvotrons must be based in an IBM PC AT chassis, but come with drivers for interfacing them to Silicon Graphics machines. The Convolvotron associated with the simulation host will be housed in an Acoustetron 3D Audio Workstation chassis (with the development option) for the proper manipulation of sounds for presentation in the VE. The manipulations performed on the development system are transferable to the other Convolvotron systems. The cost of the development system is approximately \$27,000K. The cost for the other Convolvotron systems is approximately \$15,000 plus the cost of an IBM PC AT compatible computer.

Virtual Technologies' CyberGlove (~\$10,000 each) is recommended for the VE MOUT simulator because it is the most versatile of its kind. One is required for each hand.

The tracking system one chooses is dependent on how the movement question is handled. The high degree of movement required in MOUT points to the use of the large work volume approach discussed in Section 4.2.5. A large room needs to be outfitted with a tracking system that can monitor the entire volume. The Ascension Flock of Birds will track up to 30 sensors in an eight foot cube. In addition, multiple transmitter units can be used to provide coverage over the entire room. With four participants having sensors on their head, gloves, and backs, this tracking system should be adequate. More sensors can be added if needed. The cost depends on the size of the room to be tracked.

The scenarios designed for MOUT training in the large work volume approach can be made to work if there are exercise stations that are the size of the real room. When the MOUT exercises have been completed in the station, the group leader assembles the team in the center of the room, and then, by verbal command to the computer or by gesture, the virtual scene moves to the next location. Once at the new location, the team can move about freely.

Speech recognition can be provided by a Dragon Systems dictation product. DragonDictate-30K (approximately \$5,000) has a vocabulary of 30,000 words and should be adequate for most situations encountered in MOUT. Each participant will need a speech recognizer. One development package (\$2,000) will also be needed for integration into the VE software.

Finally, the VE construction tool kit recommended is Sense8's WorldToolKit 2.0 for Silicon Graphics workstations. WorldToolKit is the industry standard and has been in use for a number of years--unlike the other commercially available tool kits.

The modeling package chosen is up to those doing the modeling. A modeler producing AutoCAD files is desirable because it is the industry standard object data format. It must be emphasized that the largest percentage of time spent in building a VE will be in the modeling phase.

Optional hardware recommended includes physiological monitoring equipment, such as the Cardiopulmonary Personal Monitor discussed in Section 4.1.9.2, the SIBIS system for evaluation of painful feedback in the MOUT learning setting, and a single sensed bodysuit for evaluation. In addition, a wireless system should be developed to handle the outputs of the glove sensors.

4.3.2 The Low-End VE Research Testbed System

The low-end system will be a single participant VE MOUT research simulator. Instead of the duplicate hardware required for the multiple participants supported in the high-end simulator, only a single HMD, image generator, etc. will be needed. The simulation host described for the high-end system and all equipment associated with it (eg. the 3D audio workstation) will remain the same.

All research issues that are not dependent on the presence of multiple participants can be explored with this system. However, team interaction is a major part of MOUT. It is recommended that this system be chosen only as a stepping-stone to the high-end system.

Chapter 5

MOUT Human Factors Issue Review

5.0 Introduction

The previous chapter has described the applicability of VE technologies for training both individual and team level MOUT tasks. This chapter identifies and examines the human factor dimension for sensory, perceptual, cognitive, motor response, instructional and training research issues associated with MOUT training tasks. It also recommends a number of research issues that should be addressed for the USMC to more fully realize the potential of VEs for MOUT training.

5.1 Task 1: Movement Through Urban Areas Outside Buildings

Urban combat requires coordination of movements both within buildings and between them. The skills identified with Task I include those that enable individuals to traverse the areas between buildings. Movement between buildings risks exposing the individual to enemy fire, and requires knowledge of the limitations of enemy surveillance, skill at firing the weapon while moving, and ability to maneuver through uneven and dangerous terrain.

5.1.1 Fundamental Skills

The behaviors, which have been identified as "fundamental skills," are generally motoric in nature, and ordinary in that they are performed in everyday life, not exclusively in MOUT. Their inclusion here results from the fact that, though the behaviors themselves are not unusual, the individual must use judgment, intelligence, and knowledge in applying them. It is not practical to consider training an individual to "exit a doorway," for instance, but it is necessary to teach awareness of the circumstances and cautions that surround exiting a doorway under urban MOUT conditions. Thus, there is a need for including these fundamental skills in MOUT training.

5.1.2 Avoid Open Areas

The tasks described in this section involve judging what areas of the environment are vulnerable to enemy fire, and, just as importantly, those areas that are not vulnerable. There are perceptual and cognitive aspects to this judgment, with training issues to consider.

Cognitively, one estimates the enemy's field of fire by identifying the enemy's position and then imagining straight lines extending from it until they strike something hard enough to stop a bullet. These trajectories are more difficult to estimate when they traverse regions of space, which cannot be seen by the individual, and are especially difficult when regions of the environment are unknown.

As the goal is to identify a covered route to a tactical Position, one should analyze regions of the environment within one's range of vision. Having identified a desirable position, one analyzes the distance of the route to it and the time it will take to move to it, especially searching for places where the enemy can have a clear shot.

From the training perspective, these tasks need to be thought about, as well as practiced. VEs can be used for training these skills, as our findings indicate that a sufficiently accurate three-dimensional view of the environment required for the analysis of cover can be modeled. The skills

can be practiced within the broader context of a war game or simulated combat in virtual environments.

5.1.3 Conduct Movement Using Cover

The marine must learn to use buildings, vehicles, and other obstacles as protection from enemy fire. These skills are motoric, but with a large cognitive component: the individual must estimate what environmental features constitute good cover. As the level of danger can be quite high, it is preferable to err toward conservative judgments.

As in the requirement for avoiding open areas, the estimate of linear vulnerability to enemy positions must be made. Further, the hardness of objects to be considered as cover must be known or estimated. Thus these tasks rely on accurate simulation of depth effects, and renderings that portray objects realistically enough to allow estimates of their hardness or suitability as cover.

5.1.4 Suppress or Obscure Enemy Fires

The suppression or obscuring of enemy fires requires a series of judgments and observations.

- The source of fires is identified.
- A plan for countering it is developed.
- The plan is implemented.
- The results are ascertained.

Because most weapons fire in a relatively straight line (with some exceptions, such as mortars and grenades), individuals should be able to identify the position of someone who is firing at them. If the enemy is firing from behind a barricade or obstacle, it is possible that linear retaliatory fire cannot be successful. Having assessed the feasibility of various countering techniques, the individual selects one and executes it.

Determining that enemy fires have indeed been suppressed requires observation of enemies who are either surrendering or wounded or dead. It may also mean that the enemy is no longer firing, has become silent. This condition frequently occurs on the battlefield. Enemies who are seen retreating might be leaving dangerous comrades behind, and the mere cessation of firing might also be a trap. Training of this skill requires some ability for moving around to a point where enemy activity can be observed; further, some probability of ambush should be programmed into a virtual world simulation.

5.1.5 Move at Night or During Periods of Reduced Visibility

Movement with reduced visibility forces the individual to rely on other senses, as well as orientation to memorized landmarks. Algorithms have been devised to represent effects of smoke, fog, and dust; these allow the user to glimpse environmental objects to varying degrees.

Nighttime conditions can be simulated as well, with a "virtual planetarium," presenting the user with an overhead sky, weather, and objects such as passing aircraft.

5.1.6 Select Routes That Will Not Mask Friendly Fires

Like conducting movement using cover, this task requires that individuals identify and know the location of weapons, and the direction of their fire; but in this case, the problem is to stay out of the line of fire so that it can continue. Again, the individual must assess the linear trajectory of weapons fire; in this task, one must also assess likely targets, which might be enemy positions or open areas requiring fires for cover, to avoid being hit by cover fires.

VEs can simulate depth of field, and present the illusion of three-dimensional space. According to our analysis, presently available technology should be capable of representing three-dimensional space accurately. However, when the perception of relationships between objects in space is a primary aspect of task performance, as it is here, VEs might not be ideally suited for training; we recommend conducting research to determine the accuracy with which distances can be represented within a HMD.

5.1.7 Cross Open Areas Such as Streets, Fields, Open Areas Between Buildings, Rapidly Under Concealment of Fires

Marines must cross spaces that offer the potential for vulnerability to enemy fire; it is thus necessary for team members to provide cover. This task is viewed from two perspectives: that of the individuals placing cover on enemy locations and that of individuals crossing the open areas. From the first perspective, the skill required is primarily cognitive. The individual covering an area with fire must be able to assess not only where enemy positions might be located, but also where the friendly fire would be most effective, for instance, blocking paths the enemy might take to escape or attack. Further, friendly fires must be timed to cover but not hit team members who are exposed. On the other hand, individuals crossing the open area must identify the best route to the objective, estimate the risks involved, and then move across the area.

These tasks offer two kinds of challenges to a VE simulation. First, while relations among objects in three-dimensional space can be represented using binocular, motion parallax, and textural depth cues, current representations may not convey depth adequately; thus, research is recommended to learn whether virtual objects should be used in training when depth perception is a primary consideration. The second challenge arises from the necessity of corporeal movement across space. Though treadmill-like platforms have been tested for some applications, VEs in which participants walk or run across some distance present a number of challenges that are described in Chapter 4.

5.1.8 Move on Roof Tops That Are Not Covered by Direct Enemy Fires

It is sometimes best for marines to travel above the plane of the street, on rooftops, concealed by eaves and other structures. This set of tasks comprises a cognitive and a motor component. The individual must cognitively assess the likelihood that enemies can fire upon particular sections of rooftops; this means visualizing the scene from various perspectives, ideally from all perspectives. This is not a trivial cognitive task, especially as it is unlikely that the entire

sphere surrounding the rooftop is visible or known. Gaps must be filled in inferentially. Motorically, the individual must walk, run, crawl, and jump across rooftops.

As with the previous task, research should be conducted to determine if the skills involved in moving on rooftops are suited to VE training with currently available equipment and software. The perception of three-dimensional space is primary, and movement of the body across space presents problems as well.

5.1.9 Select Subsequent Positions Before Moving

There are usually several routes from one point to another, each with its advantages and its disadvantages. The marine who intends to traverse a region must compare the risks and benefits of the various paths before moving. This judgment or decision-making task requires the ability to see or infer the relative positions of enemies and the proposed routes. Thus, it presents the previously mentioned difficulties regarding three-dimensional presentation. VE training allows “hints” and pointers to be displayed for new trainees, illuminating or indicating preferred positions for movement, as well as x-ray views of occulted locations and objects.

5.1.10 Move Around Corner of Building

The marine can move by pushing the body with the toes while lying flat on the ground. This basic motoric skill can be practiced within the context of a VE urban terrain, with SIBIS or other consequences for failure to keep a low profile. When the user is attached to computing equipment by wires and apparatus, movement such a crawling movement may be impeded; if illusory movement is simulated through use of a 3-D mouse, however, the perceptual and cognitive aspects of this task could be trained in a VE without special encumbrances or difficulties.

5.1.11 Moving Past First Story Windows

This task, and the following three, requires moving past an opening without exposing oneself to enemy fire. Though corporeal movement is trained in this subtask, it is not movement across large areas; the skill of passing a first-story window could be trained in a VE simulation. The software scans for input from the participant, which is “visible” from within the window, then shoots at individuals who are seen. The SIBIS technique described in Chapter 4 can be used to inflict pain and to simulate being hit by weapon fire, or simple visual or auditory feedback can be used to signal that the participant has been shot.

5.1.12 Moving Past Basement Windows

This task presents the same challenges as the previous one, and problems can be addressed with the same techniques. Individuals must ensure that they are not visible to basement dwellers for a length of time, which would allow them to become targets of weapons fire.

5.1.13 Crossing a Fence or Wall

This task requires reconnaissance and estimation of risks, coordinated with motoric behavior. In our analysis, the skill of crossing a fence or wall can be trained in a VE, and it can be greatly improved when the VE is augmented with workarounds. A virtual fence or wall can

obviously not be climbed, but a workaround such as a sawhorse or other solid device can be integrated into the VE to allow training. Booby traps should be randomly programmed into the VE with some probability, and their locations varied between trials. The software can scan for "visible" targets, and fire at individuals who present too high a profile or take too long: it is probably best to program a slight lag, simulating the enemy's reflex latency, so that individuals who cross the barrier within some criterion will not be penalized. Those who do not meet the criterion, however, should be penalized either through SIBIS simulated gunshot or some visual or auditory mode.

5.1.14 Passing or Exiting Doorways

As with previous tasks, the individual can learn to keep a discreet profile and more quickly while passing openings of buildings. Consequences can be administered when criterion visibility thresholds are surpassed or the movement is too slow, including painful SIBIS "gunshot" and visual/auditory weapons and explosives effects.

5.1.15 Movement in Streets

Moving in streets potentially exposes the marine to danger from enemy fires. If this vulnerable movement is necessary, the individual should ensure that cover is adequate and/or that concealment through use of smoke and suppressive fires is available. Training for movement in streets ideally comprises a classroom component and a practice component. Individuals should learn in a pedagogical setting that movements should be conducted inside buildings whenever possible, how to select routes, which provide concealment, and how to use smoke and suppressive fires for cover when necessary. The actual movement through streets can be practiced in a realistic VE urban setting populated with enemies and locals.

Movement across open spaces such as streets can be accomplished by several approaches. The room within which VE activities take place could be large enough, and electrical tethers long enough, to allow users to traverse wide spaces. A second approach is to develop a treadmill-type apparatus to impart the illusion of movement with the individual remaining in place. Finally, users can "fly" from one location to another. Advantages and disadvantages of each technique are discussed in Chapter 4.

5.1.16 Movement Across Open Areas

As in the previous task, movement across open areas often results in vulnerability to enemy fire, with important negative consequences for failure to make the correct decision based on minimal information. Movement can be practiced in a realistic virtual urban environment. Aspects and methods of simulating movement of the participant should be researched thoroughly, and an appropriate technique selected for particular training tasks.

5.1.17 Select, Occupy, and Use a Hasty Firing Position During Movement

The individual should be trained to make use of corners of buildings, walls, rubble, overturned vehicles, and rooftops for hasty firing. An important aspect of this skill is the necessity of the objects to support the weight of the participant and/or his weapon; therefore, in order to simulate these situations, workarounds can be used with virtual reality techniques. Workarounds

introduced into the VE should be irregularly placed, and should vary from one training session to the next. Further, in a realistic urban context, enemies should appear from varying locations, to force trainees to learn to locate them and to fire in all directions.

5.1.18 Firing the Individual Weapon During Movement

As discussed in Chapter 4, the weight and kick of a weapon almost necessitates training with nonvirtual weapons. Of course, workarounds, in the form of real weapons firing blanks, can be integrated into a VE paradigm, and the individual can practice firing at virtual targets. The limitation is that the HMD covers the eye, and thus altered weapons may have to be devised, with virtual sights. Despite this limitation, this task is suitable for VE training; individuals can practice firing in various positions and situations.

5.2 Task 2: Human Factor Issues in Entering Buildings

5.2.1 General

The general training requirements include cognitive, perceptual, and motor skills, which may be trained simultaneously or separately. First, an analytic comparison of possible points of entry into a building requires some knowledge of the advantages and disadvantages of various entryways. Knowledge to inform this judgment can be imparted through classroom training and/or text, as well as experience with exercises and mock combat. Individuals must learn how to check holes and windows, etc. for booby traps, as well as how to place booby traps.

VEs can be used to train individuals in these skills; individuals could search for booby traps, or select an entry point, and virtual consequences could be administered. For instance, failure to search for a booby trap in a likely spot, could result in a computer generated warning, perhaps an illuminated “x-ray” view of the concealed booby trap, followed by a simulated explosion, or whatever effect the trap might produce. For more extremely critical work, it may be desirable to inflict pain (via SIBIS or other method) to simulate the consequences of weapon fire and explosion.

We recommend the use of VE practice in the early stages of training, to introduce individuals to unfamiliar skills; further training should subsequently take place in a real-world setting; and finally, refresher training can be conducted in VEs to maintain skills.

5.2.2 Fundamental Skills for Entry Techniques

Trainees must learn to enter upper levels using grappling hooks, scaling walls, and entering windows, descend using rappel techniques, and enter lower levels using various one- and two-person lifts. Each of these tasks presents a particular problem for training in virtual environments, which is that an object or person must support the individual's weight during task execution. Of course, virtual objects cannot support weight, so a virtual training environment should be augmented with workarounds. In a workaround paradigm, the individual would climb a real rope, or be lifted by a real person, while viewing the VR representation of an environment.

5.2.3 Entering Upper Floors

The marine must learn to use a grappling hook for entering upper-story windows. Skills involved include selection and use of materials, as well as avoidance of exposure while passing lower level windows, clearing rooms, and covering other marines, once the upper level is attained. Virtual grappling hooks present no resistance to the user, and training in a pure VE would probably not be especially effective; similarly, one can not climb up a virtual rope. We have recommended, in Chapter 3, a program of research aimed at development of a repertoire of workarounds to be adapted for tasks such as these that require physical force. In this particular case, actual objects resembling grappling hooks in weight and feel could be introduced into the VE, with real ropes attached. The flight of the hook and rope could be represented in the H:N4D, and the user could "climb up" the rope through an illusion created by counterweights, pulleys, and a representation of ascent in the HMD.

5.2.4 Entering Middle Floors

Middle floors are entered through performance of various one- and two-person lifts and pulls, and through sling lifts. The one- and two-person lifts require cooperative effort, as well as support of the lifted individual's weight; as such these introduce a challenge beyond that of the previous task. Coordinated interpersonal effort can be programmed through multiuser simulations, or through artificially intelligent programmed virtual actors, with workarounds to do the actual lifting.

5.2.5 Entering Ground Level Floors

Ground floors may be entered through doors or through openings, which are created with explosives and artillery. The individual must acquire skills with explosives, and must learn to be sensitive to potential booby traps. The cognitive aspects of training ground-level entry can be addressed through VE training techniques. Individuals could practice searching for booby traps, placing explosives, and taking cover from the blast.

5.3 Task 3: Clearing Rooms

5.3.1 Introduction

Skills involved in clearing rooms are multifaceted, with motor, cognitive, perceptual, and social components. Individuals must learn to prepare and detonate explosives, coordinate movements among team members, control volatile interpersonal situations when rooms are occupied, and move among and between locations. VE simulation required for training must represent the world dynamically, with unexpected events, difficult-to-identify people, randomly placed booby traps, and frequently highly emotional behavior of others. As such, these skills can be trained effectively in a VE with appropriate planning, expertise, and technological infrastructure.

5.3.2 General

The general techniques involved depend on specific skills described in subsequent sections. Basics such as gaining surprise, clearing the room quickly, and overwhelming the occupants of the room, however, should be communicated to trainees first as general principles. This educational task can be accomplished efficiently through text or pedagogical techniques, followed up with practice in a VE. As these general tasks focus on elements of surprise and seizing the psychological advantage, these skills can be practiced in a realistic multiuser combat simulation, with users divided into teams of friendlies and enemies.

5.3.3 Assign Sectors of Fire

This team skill requires coordination among individuals, under direction of a leader. The simulation of interactions among team members requires programming of human behaviors, which will inevitably be addressed. Team members' interactions can be simulated either by representation of multiple participants, or by simulation of behavior of others. Though the first option requires more powerful hardware, it is currently the more feasible of the two. On the other hand, the ability to realistically simulate human behavior would represent a quantum increase in the capability of virtual reality as a tool for applied and basic research, and as such we recommend continued investment in research toward virtual modeling of human behavior.

5.3.4 Eliminate the Threat

The essence of eliminating the threat is judging whether a person is an enemy or not. Often in urban warfare, individuals wear no uniform or other signs to clearly identify them as friend or foe; in fact, individuals may switch sides from day to day or even work for both sides simultaneously. A recent episode in Somalia, for instance, resulted in the death of U.N. forces at the hands of women who pulled weapons from under their dresses. Thus, an individual must learn to observe cues to indicate the presence of danger.

Training of these skills in VEs requires the programming of realistic virtual people. As the Somalian example indicates, indigenous clothing may hide or provide cues, and thus individuals to be assigned to a particular region should have experience interacting with natives.

The programming of realistic human behavior presents several interesting research issues for research. The present problem requires depictions that are realistic enough to allow observation of bulges indicating concealed weapons. It also requires animation of representations in a way that accurately simulates a culture's norms. As social interaction such as that found among team members need not be simulated in the representation of locals, the techniques required are concrete research topics that can and will be addressed in the near future. The representation of realistic human behavior is a necessary step in the evolution of virtual environments.

5.3.5 Control the Situation and Personnel

Individuals entering a room must determine whether the occupants are living or dead, command authority over the living, and search the room for threats. Trainees should learn to speak concisely, and in a loud commanding voice. Authoritative speech can be trained in a VE, with the

participant speaking to others or to animated virtual persons. This might seem unnatural to the participant, yet it could be hypothesized that the training would transfer efficiently to the combat situation, where persons in the room are not to be treated with personal consideration. Programming would be challenging, if virtual persons were programmed to respond to the participant's tone of voice (prosody) and/or content of commands. Responding to the challenge, however, would advance VE technology greatly, and open up new domains of training paradigms maximizing usage of voice recognition and natural language knowledge.

Vitality of bodies is determined by the "eye thump" technique, which requires the resistance of a solid object, and consequently can be simulated with workarounds. Further tasks involve searching the room and live persons for booby traps and other threats, and searching the dead. These skills could be trained using VE methods.

5.3.6 Secure and Evacuate Personnel or Equipment While Maintaining Rear Security

The difficulty with training for this skill lies in the probability that it will not be recalled at the time it is relevant. That is, an individual immersed in a combat situation may lose sight of the priorities that originated the mission. As Zajonc (1965) has pointed out, individuals under stress tend to emit the dominant, or best learned, response. Therefore, the behavior should be practiced in the context of a virtual urban terrain until it becomes automatic.

5.3.7 Organization

As in the previous item, individuals need to learn the organization of the team, and this learning needs to be available in stressful conditions. The tasks of covering while a search team clears a building can be practiced in exercises and mock combat in a virtual urban environment.

5.3.8 Clearing a Room

This task breaks down into three sections or skills. First, individuals must cook off a grenade and throw it into a room; next they enter the room, following a formal procedure with behavioral and social components; and finally they mark and leave the room, to repeat the procedure elsewhere.

Entering the room according to the prescribed procedure, calling out "Clear" or "Coming in, last name," should be practiced in team exercises. Leaving the room and searching the entire building can be trained in a VE, with enemies and booby traps hidden in random positions in the building.

5.3.9 Use C-4, Claymore Mines, and TNT Demolitions to Gain Access to Rooms

The simulation of explosions presents a number of challenges to VR programming. The trainee must master two basic tasks: placing the explosive for maximum effectiveness and moving out of the way of flying debris. To model these tasks, it is necessary to plot the chaotic paths of debris, which are affected by such parameters as hardness of walls, contents of the room and their hardness, type and amount of explosive, angles of reflection, and trajectory arc as a function of

gravity and mass of object. In order for training to be effective, these parameters must be calculated accurately.

Fortunately, there is a temporal lag between the moment of the placement of explosives and their detonation. Thus, the CPU can be calculating debris trajectories in the quiet time before detonation. The algorithms required for this programming have not yet been developed, though they will be crucial for modeling of many military situations. While VEs may not meet these challenges at this time, it should be possible to develop this capability in a relatively short space of time should this be desirable.

5.3.10 Clear Rooms with Closed Doors

This is essentially the same as clearing a room, differing mainly in the argument against using doorknobs to open doors. Individuals open the door with explosives, which, as mentioned in the previous task, presents rich challenges to VR programming. The strength of the explosion is a function of the mass of the explosive, and the flight of debris, including destruction within the room and without, is a function of the strength of the explosion and the position of the explosive relative to other objects. From a training standpoint, the individual must learn to use adequate force for the job, without risking injury to self or team members. This requires the marine to understand the effects of various amounts and placements of explosives. VEs present a number of advantages for learning dangerous, destructive, or potentially injurious skills without risk to the individual.

5.3.11 Search and Clear Basements

This skill includes room clearing and checking for booby traps, plus checking for sewers and tunnels. It relies heavily on high-resolution graphics for the inspection of features of the room. This task can be trained in a VE, with randomly concealed openings to sewers and basements, and randomly occurring consequences for missing them.

5.3.12 Avoid Hallways If Possible; If Movement Along Hallways is Necessary, Move Along the Side of Walls as Quickly as Possible to Get Out of the Hallway

In avoiding hallways, the individual will frequently create openings between floors of a building, using explosives or heavy weapons. The individual also must check for booby traps and presence of enemies. These skills can be practiced in a VE combat simulation, with negative consequences for presenting too visible a profile. Enemy fire should erupt from random, unexpected locations both inside and outside the building.

5.3.13 Move Between Floors

This task involves checking stairways for danger and, if necessary, using explosives or heavy weapons to open holes in ceilings, floors, and walls. The issues here have been addressed previously: the simulation should present random threats, and debris from explosions needs to be modeled accurately. Training can take place in a VE context with workarounds.

5.3.14 Mark the Building Once it Has Been Completely Secured and Announce "All Secure"

This task can be practiced in VE and real-world exercises, focusing experience on realistic combat simulations. Frequent practice is the best guarantee that the individual will remember this rule in a combat situation.

5.3.15 Reorganize the Force

After securing rooms, the team must replenish and/or redistribute ammunition and deal with the wounded. This skill can be practiced in team exercises using a multiuser VE paradigm. Field triage is a particularly difficult skill to acquire, as examples of real wounds normally do not exist for training exercises. Virtual reality, however, offers the possibility of demonstrating different types of wounds, to include life threatening wounds, realistically. Present technology will not enable the marine to treat wounds, but the simulation could elicit voice input, with the individual telling what to do, given the visual display of a particular type of wound.

5.4 Task 4: Establishing of Defensive Positions

5.4.1 Fundamental Skills

The fundamental skills include selecting and establishing weapon positions with fortification, communications, fire prevention, and secure routes to other locations. The skills are concerned with selection and preparation of defensive positions; thus cognitive and motor skills are involved. These should be taught separately and then combined.

Trainees should receive text and pedagogical training in factors that contribute to good personal, tank/APC, ATGM, and sniper positions. They should also be instructed to secure and fortify buildings, take fire prevention measures, and establish communications.

After techniques have been described, through text, video, or classroom instruction, individuals should practice these skills. Selection of good positions can be performed in VE simulations, but preparation, involving the manipulation of building materials, can best be trained in real-world environments. Finally, the skills should be practiced in a combat simulation.

5.4.2 Identify and Take Up Hasty Firing Positions

Hasty firing positions can be practiced within the context of virtual environments. Individuals should have already learned to fire their weapons from all positions, and that skill can transfer to the use of workaround weapons, with approximately the weight and shape of real weapons but trackable by the system and cut away to allow placement of the weapon near the face without dislodging or damaging the HMD.

5.4.3 Prepare a Defensive Fighting Position

The individual is required to move objects, dig, and knock holes in walls; these skills are suited for VE simulation with workarounds. Individuals can be instructed to recognize features

that make a good defensive fighting position, and then practice these forceful behaviors in virtual exercises.

5.4.4 Establish Communications With Other Members of the Defensive Team Via Wire or Radio

Individuals should receive instruction in operating wire and radio communications; this skill can then be embedded in the context of real and virtual combat exercises. Workaround communication equipment could be developed to simulate the operation of the apparatus.

5.4.5 Establish Defensive Fires for Different Weapons for Assigned Sectors

The skills involved in setting up various weapons (machine-guns, anti-armor weapons, etc.) contain cognitive and motor components. The knowledge involved in selecting a position for the weapon and clearing areas around it, if necessary, can best be conveyed in a pedagogical context, either a classroom or a manual to be studied. Motor skills here are of two types: first, individuals must know how to operate the various weapons, and second, they must learn to transport and prepare the weapons for firing. The first of these skills can be trained using classroom or text methods, and the second should be practiced in VE exercises with workaround weapons.

5.4.6 Employ Integrated Obstacles, Barriers, and Defensive Fires

This task can be simulated in a VE paradigm, with the development of weighty workarounds. The tracking system must be able to detect the positions of these obstacles, especially if they are to be moved, and they must be designed to prevent entanglement with electronic apparatus.

5.5 Task 5: Human Factors Issues That are Common to All MOUT Tasks

5.5.1 Introduction

The skills, which are common to all MOUT tasks, should be trained separately until trainees have a fundamental understanding of their execution, followed by practice in the context of a combat simulation in a virtual urban terrain. Workarounds devised for particular tasks should be versatile, usable in isolated skill training as well as the generalized combat situation.

5.5.2 Throw Grenades

Throwing a grenade accurately requires coordination of sensory and motoric variables: the weight of the grenade must be controlled by the force of the arm movement in such a way that the object is propelled toward its target. However, cooking off a grenade can be trained in VE with workarounds. In fact, VE should be very useful in demonstrating the results of mistiming the release; the grenade can explode in the trainee's hand or the enemy avoids it if the grenade is thrown too early.

Training in throwing hand grenades relies on proprioceptive feedback, and probably requires the individual to throw an actual object, which resembles a grenade in weight and size.

The actual act of throwing into windows is most efficiently practiced with an actual object and a real window.

5.5.3 Use Camouflage Techniques

The detection of a foreground object against a background may be accomplished by use of any of several visual cues. The goal of camouflage techniques is to "trick" the perceiver into seeing background only. Representations of individuals in a VE, however, are arbitrary: thus practice of this skill in a real-world situation and not in a VE is recommended.

5.5.4 Booby Traps

One of the best ways to learn to detect and avoid booby traps is to learn how to make them. While the enemy may possess methods we do not know about, it is prudent to assume that they know everything we know. Thus, the trainee who learns how to make booby traps also learns how to discover them.

The detection of booby traps can be modeled in VEs, with the participant being required to search for different possible installations. Consequences, either SIBIS shocks, depiction of booby trap consequences, or suprealistic indicators (buzzers, etc.) can signal errors.

5.5.5 Mines

As with booby traps, one of the best ways to learn to detect and avoid mines is to learn how to emplace them. Detection of mines can be simulated in VEs, especially when graphic resolution is very fine. X-ray views and indicators can assist the trainee in identification of cues to the presence of mines. Again, consequences should be administered when mines are not discovered.

5.5.6 Use Obstacles

The use of obstacles sometimes implies the manipulation of large objects, which can provide cover and protection. The force exerted by objects against the individual is difficult to simulate, and, in this case, the result might not justify preferring VE over real-world methods. On the other hand, when obstacles are already in place, such as walls, vehicles, rubble, etc., the individual's task is to protect himself from enemy fire; this behavior can be modeled. The obstacles are represented graphically, and if the participant exposes himself to enemy fire he will experience consequences such as SIBIS shock, portrayal of fires, or suprealistic indicators of error.

5.5.7 Talk With Other Members of the Team Involved in MOUT

Personal interaction is an important factor for training team skills. Multi-user VEs can be used to simulate team behaviors, and conversations can be telecommunicated among members, even when they are actually remote from one another. The requirements for multi-participant simulations are complex and will be addressed in the near future. The issue of personal interaction and communication is discussed in Chapter 3.

Chapter 6

Research Summary and Conclusions

6.0 Introduction

This research concludes that VEs can be used effectively for training a number of MOUT skills. There remain, however, a number of unanswered questions regarding the optimization of VE training for MOUT applications. This chapter will summarize our recommendations for development of a virtual MOUT training environment, including suggestions for research into important factors affecting training success.

The analysis finds, first, that many MOUT tasks can be successfully modeled and trained in VEs. VE training will be especially useful in the initial stages of training, where marines can become familiar with complex, dangerous, and/or expensive equipment and situations, and also in the maintenance or refreshing of skills after they have been learned. For some skills, it appears preferable to conduct most of the training in real-world situations, after a VE "initiation," with follow-up practice in a VE paradigm to maintain skills and to prevent skill decay.

A second finding involves the identification of important areas for future research. Given the newness of VE as a training context, there is uncertainty as to the parameters of effectiveness of simulated environments for training: in particular, there is a need to determine factors that influence the transfer of training from the VE to real-world performance of the task.

The present discussion will begin with a description of a recommended paradigm for VE training, followed by recommendations for research into factors affecting the outcomes of VE training applications. Recommendations will be kept general to suggest a range of possible future activities in the development of VE paradigms for MOUT training.

6.1 Recommended MOUT Training Paradigm

An obvious first step in VE MOUT training entails creation of a virtual urban environment. Models of cities presently exist that may be adaptable to MOUT requirements. The urban environment should be consistent with USMC MOUT specifications to facilitate the transferring of skills. All buildings should be completely modeled, interior as well as exterior, allowing entry and movement within structures as well as on and between them. Current VE technologies are sufficient for creating the required terrain model of urban areas.

6.1.1 Population

Warfare around anthropogenic obstacles differs from warfare in a forest, jungle, or other natural terrain largely as a function of the presence of people, including friends, foes, those of unknown allegiance, and innocent bystanders. It is, therefore, necessary to populate the urban environment. As described in a later section of this chapter, artificially intelligent virtual people can be selected from a range of prototypes, from simple preprogrammed repetitive actions to agents with autonomy and intelligence. The environment may also be populated with representations of other participants in a multiuser paradigm.

Friends: Team members can be programmed virtual persons with some natural language processing ability, and capabilities for responding to verbal communications with the

participant. Task definitions will be straightforward enough to allow fairly realistic behavior to be programmed without excessive use of computational resources.

Coordinated behaviors between the participant and a virtual actor, such as carrying the wounded, two-person lifts, etc., are possible to program, but relatively difficult, and the difficulty is probably not justified when one considers the relative low costs of real-world practice.

Foes: Enemy soldiers can be animated in an urban scene, running across the street, shooting from random windows and rooftops, and can be discovered inside randomly selected buildings. One important aspect of representation of enemies is surprise, meaning that the probability of an individual appearing in a predictable location should be unknown to the participant. As in actual combat situations, enemies should ambush from unlikely spots, and fire from corners and windows that the participant may not anticipate. It is important that enemy soldiers not appear in the same sequences and locations, but appear at random from logical locations.

Undeterminable Allegiance: Urban warfare presents many circumstances where the allegiance of individuals cannot be determined. Enemies may wear friendly uniforms, or civilian dress, and may be women, children, or the elderly. Virtual modeling of such situations requires a rather high degree of pictorial realism, as the participant must visually sort subtle cues, such as nonverbal behaviors, carrying a weapon, bulges under clothing, etc. It will be worthwhile to train soldiers in local customs, especially as these present opportunities to assess the friendliness of strangers. In summary, a VE town or village can contain a realistic number of natives whose loyalty is not obvious.

Innocent Bystanders: Urban warfare is often played out in busy streets and crowded areas, where most of the individuals present are not participants in bellicose activities. The behavior of these people may largely be programmed in routine, repetitious or fixed patterns, as a computer resources. In a combat situation, however, with gunfire, explosions, and soldiers moving about, the behavior of bystanders should be programmed to accelerate as individuals take protective actions.

6.1.2 Workarounds

Many MOUT skills require the presence of solid objects, either to support the trainee or to otherwise apply force against him. It must be acknowledged from the beginning that some tasks can simply not be performed in a VE: for example, one can not climb a virtual rope. However, many of the limitations of virtual reality can be overcome with the ingenious integration of actual objects with virtual ones; for the present discussion, the actual objects are referred to as "workarounds."

The development of a repertoire for workarounds comprises a major effort in itself. While problems can be roughly categorized into types, in a more pragmatic sense each problem will require a specific solution. A few of these will be suggested next, along with a discussion of limitations, which cannot be overcome through this approach. A recommended research program to address workaround issues is presented more fully in the third section of this chapter.

Rope Climbing: Some MOUT tasks require the participant to climb up or down a rope. While a virtual rope obviously cannot support a person's weight, we can conceptualize a straightforward pulley arrangement, whereby a real rope is suspended over the VE area. The rope is counterbalanced over a pulley, so the user can pull up to a few feet off the floor; tugging on the rope after that results in the rope looping through the pulley at the same rate that the individual "climbs." The representation in the HMD is simultaneously shown from a higher perspective, again synchronized with pulls on the rope. Eventually the individual will reach the second story (or the ground, if the marine is descending) and the floor of the laboratory will simulate the new level.

Rifles: The marine must be constantly prepared to use a weapon. While a weightless representation of a rifle could be displayed, there are reasons for using a workaround rifle. It is not practical to use real rifles, though the heft and kick would be maximally realistic. One reason is that the HMD will conflict with the sights of the gun. Therefore, we recommend development of workaround weapons.

The ideal workaround rifle will be the weight and size of a real weapon, so marines become accustomed to carrying and firing it. Our recommendation, however, is to devise model weapons with a section cut out so the individual can hold the weapon up to the eye correctly. The system will need to be able to track the position of such workaround rifles; thus, depending on the type of tracking system used, it may be necessary to attach some apparatus to the device. The most important and difficult problem to overcome with this simulated rifle will be the accuracy of tracking. It is not acceptable to train marines on weapons with inaccurate sights or biased aim of any kind.

Windowsills and Walls: Sawhorse-type apparatus could be installed in the laboratory to model solid objects to be climbed upon and over. Again, a primary task is to accurately "track" the positions of such objects in the virtual representation. A more daunting problem arises from the difficulty of moving with an umbilical cord of wires connecting one to the maternal computing hardware. Further, even if wireless equipment were used, trainees would be limited in the amount of vigorous movement (jumping, running, and tumbling) they could engage in without damage to sensitive and expensive equipment. Still, workaround windowsills, ledges, and walls are possible in modeling those MOUT situations, which do not require strenuous or jolting movement. However, the cost of developing an effectively integrated system of windowsills, ledges, and walls may be more than practicing on real structures.

Movement of the Participant Through Space: As described in Chapter 4, the ability to track an individual's movement through the work volume varies widely among tracking techniques, systems, and VR models. Some researchers have reported good results with participants walking in place, with the direction of apparent movement determined by head direction or other cue. It is recommended that, if stationary walking is used to simulate movement through space, direction should be determined by sensors in the soles of the shoes, rather than HMD position or glove position, to enable the individual to look around, operate machinery, etc., while moving, as in actual combat situations. As direction of walking is determined by the direction of the forces applied between the foot and the ground, this method of tracking direction seems to offer the greatest verisimilitude and the least interference with other behaviors.

As a way around the problem, VR paradigms frequently use a “flying mouse” or three-dimensional trackball to permit the user to fly through the virtual environment. While this technique does enable the trainee to travel from one virtual location to another, it does not provide practice in traversing areas, which in combat can be a major challenge.

Other workarounds for a wider range of movement might depend on adaptation of the treadmill. One method might use a 360 degree treadmill apparatus, which is capable of turning in response to pressure from the individual's feet; this device, which is currently unavailable, could represent the illusion of movement to the stationary participant.

Painful Consequences: By definition, a VE is one in which actual consequences do not exist. Combat, on the other hand, is a situation in which the consequences can be very high indeed, with life, injury, or death often balancing on a subtle distinction or hasty decision. Soldiers in combat may be hit by gunfire, struck by shrapnel from grenades and mines, may be hit with pistols and other blunt instruments; a VE simulation, which is able to inflict pain, will not only appear more phenomenologically realistic, but will result in improved transfer of training, and trainees will learn to really avoid certain risky situations. Scientists of the Johns Hopkins Applied Physics Laboratory have devised an electroshock device (SIBIS), which delivers a painful but nonlethal voltage to the user (see Chapter 4). Such consequences may be useful for training tasks like walking past windows or climbing over walls, where errors in combat could result in painful consequences.

6.2 MOUT Tasks That are Appropriate for VE Training

This report subcategorizes MOUT skills into five discrete classes of tasks. The following section will describe the usefulness of VE training for the various tasks.

6.2.1 Task 1: Training Requirements that Must be Fulfilled for Movement Through Urban Areas Outside of Buildings

- 1. 1: *Fundamental skills (generally motor).*
- 1.2: *Avoiding open areas.*
- 1.3: *Conducting movement under cover.*
- 1.6: *Selecting routes that will not mask friendly fires.*
- 1.7: *Crossing open areas such as streets, fields, open areas between buildings, etc., rapidly under concealment of fires.*
- 1.8: *Moving on rooftops that are not covered by direct enemy fires.*
- 1.15: *Moving in streets.*
- 1.16: *Moving across open areas.*

The Task 1 skills listed above involve moving across distances that are large relative to VE workspace. The usefulness of VE modeling of these situations depends on the technique that is chosen for movement through space. Three basic approaches to the problem have been identified. First, simulations can be conducted in a large room, with a tracking system that can follow the participant's movements over a relatively expansive spatial range. This approach would provide the advantage of allowing realistic movement, along with the disadvantage of decreased portability. A second, middle-of-the-road, approach is to use treadmills or similar devices to simulate movement across areas. Treadmills may convey the impression of movement, but the user is restricted to movement in one direction only, or else a "steering" mechanism must be developed to enhance the illusion of ambulation. Finally, participants can "fly" through space using a mouse or joystick; this approach maximizes the amount of virtual space that can be covered, but minimizes the realism of movement.

Additional research should be conducted into the effectiveness of each of these approaches. At one extreme, it may be important for the marine to experience actual motoric movement across a space; on the other, it may be sufficient to click on an icon, for example, to jump through a window, causing the HMD representation to display the moving environment while the participant sits in a chair.

1.4: Suppress or obscure enemy fires.

This skill can be trained in a VE resembling current video games, with enemies located in various positions in the virtual world. The simulation should include snipers and other forms of enemy fires, and should require the participant to discover the source of the fires, counter them, and finally to ascertain that the task is completed. Though modeling of simple enemy behaviors is not especially problematic, it is important that the script be versatile, and that enemies appear in new locations with unknown probability.

1.9: Select subsequent positions before moving.

Whereas "moving" itself is difficult in a VE simulation, it should be possible and practical to have the participant select a subsequent position. Using a hand held mouse or other device, the individual could test various paths to the goal position, discovering whether these paths are safe or not. SIBIS consequences could be administered to make the simulation more realistic and more meaningful to the trainee.

1.10: Move around corner of building.

1.11: Moving past first story windows.

1.12: Moving past basement windows.

1.14: Passing or exiting doorways.

1.17: Select, occupy, and use a hasty firing position during movement.

These are movements that can be trained in a VE, as the user does not cover a prohibitive amount of ground. As in 1.9, enemy presence should be unpredictable, and SIBIS

consequences can enhance the realism of the training scenario. It is recommended that these skills be integrated into the virtual world.

1.13: Crossing a fence or wall.

As stated above, this technique cannot be modeled in a virtual environment, but implementation of workarounds can circumvent problems. Trainees should learn to cross fences and walls within the context of urban combat situations. It is recommended that initial training be conducted in a real environment. Refresher training can be conducted in a VE simulation.

1.18: Firing the individual weapon during movement.

As stated elsewhere, workaround weapons need to be developed. These should weigh approximately the same as real weapons, and the experience of firing the weapon should also be approximated.

1: Summary.

Many Task 1 skills can be integrated into a virtual training environment. In general, these skills can be practiced in context until they become automatic. When the trainee encounters difficulty with a particular task, however, that task can be isolated and practiced on its own.

6.2.2 Task 2: Training Requirements for Entering Buildings

The skills in Task 2 comprise techniques for entering buildings. The trainee not only learns to climb and enter windows, but must learn to search for booby traps, place explosives, “cook off” hand grenades, etc.

Such skills as rappeling, scaling walls, and using one- and two-person lifts are not well adapted for VE training. The integration of workarounds into the virtual world, as described above, is recommended for simulating these situations. Thus, the proposed virtual would include the following workarounds:

- Rope/pulley arrangement for rappeling, grappling hook, and other skills involving rope climbing.
- Simulated hand grenade to be thrown in VE workspace.
- “Sawhorse” windowsills, fences, and walls.

Searching for booby traps makes up an important aspect of Task 2. Simulated booby traps should be realistic and unpredictable. Individuals who overlook booby traps can be offered an array of consequences, for example, the “novice” consequence could include an x-ray view of the hidden booby trap; the “intermediate” consequence could be a simulated explosion or booby trap effect; and the “expert” who overlooks a booby trap could experience a virtual detonation plus SIBIS consequences.

Placement of explosives to create entry openings in walls and ceilings presents an unusual opportunity for VE training. Assuming that a realistic explosion algorithm is devised, trainees could gain very valuable experience in preparing and detonating TNT and plastic explosives, without danger to persons or damage to structures.

6.2.3 Task 3: Training Requirements for Clearing Rooms

Task 3 includes skills for clearing rooms. These are largely team-coordination skills, which offer particular challenges to the VE simulator. Two general approaches can be considered in simulating interpersonal interactions involving the participant and others. First, a multiuser VE can allow realistic interactions, as real people interact with representations of other real people in real time. The burden in this case is on computation, with each person requiring a unique image to be generated, which in turn requires computing the positions and movements of all other users.

The second approach is programming knowledge based virtual actors to interact with the user. While this approach offers more flexibility of scheduling and reduces dependence of each trainee's regimen on behaviors of others, knowledge based algorithms have not yet acquired a level of sophistication, which allows simulation of complex, realistic human interaction.

Multiuser systems are recommended for training in team activities. Integration of multiple participants will be rather intensive in its use of computer resources, but will offer realistic interaction and ability to coordinate movements of individuals.

6.2.4 Task 4: Training Requirements for Establishing Defensive Positions

The establishment of defensive positions can be practiced in virtual environments with workarounds. Again, it is suggested that VE be used in the introductory stages of training, to familiarize marines with fundamentals; intermediate training should include field exercises and real-world practice; finally, after skills have been well trained, the VEs can be employed to keep skills sharp and fresh. For Task 4 skills, the simulation needs to include representations of such large objects as tanks, APCs, and large obstacles such as overturned vehicles, as well as workarounds to simulate communications equipment.

6.2.5 Task 5: Fundamental Skills That are Common to All MOUT Tasks

The "fundamental skills that are common to all MOUT tasks" can be embedded in training using VE, and should be practiced in real-world exercises as well. Some basic skills are not good candidates for VE training as it presently exists, including:

5.1: *Throwing grenades.*

5.2: *Camouflage techniques.*

5.5: *Using obstacles.*

Though throwing grenades can be practiced in a VE with workarounds, a computer simulation of the flight of an object through space coupled with a workaround imitating a grenade will be more expensive and less effective than practice with actual dummy grenades, which can be

thrown at real targets. The effectiveness of camouflage depends on the accuracy of color and textural matches to the background, and thus only a perfectly realistic graphical representation in a multiuser VE would be practical for learning this skill. Finally, though workarounds could be used to simulate the presence of obstacles, the difficulties of tracking these objects as they are moved around by participants, and of keeping workarounds disentangled from wires and other apparatus, suggests that this skill is not amenable to VE training.

Our analysis has identified the following Task 5 skills as being suitable for VE training.

5.3: *Booby traps.*

5.4: *Mines.*

Assuming that a good explosion algorithm is constructed, a VE simulation is suitable for practice in just this kind of dangerous skill. Trainees can gain valuable insights by handling booby traps and mines in a virtual world. The user can learn what to expect of various types of booby traps, and can learn what to look for and what to expect from various kinds of mines, all without the fear of injury to self, others, or equipment.

5.7: *Talking with other members of the team involved in MOUT.*

Difficulties with this kind of team-coordination skill are discussed above, in connection with Task 3. Implementation of a multiuser VE, allowing individuals to communicate with one another in real time is suggested.

6.3 Research Issues

Training in VEs is in an early stage of development, with many issues unresolved, many problems unsolved, and many questions as yet unasked. This section addresses research issues that should be pursued in the development of training technologies. Tasks are categorized by the requirements that affect the kind of VE training that is possible; a final section will summarize the research issues involved and suggest future research directions.

6.3.1 Physical Objects in the Actual Environment (Workarounds)

Whereas some training tasks can be simulated by the simple presentation of visual and auditory stimuli, other tasks may require the presence of actual objects, or "workarounds," in the room with the participant. This class of training tasks can be further subdivided into (1) objects that must apply force feedback to the participant, and (2) objects requiring tactile feedback. Objects in the environment that press upon the individual, or, more frequently, upon which the participant will press, will be included in the force feedback category, and objects with texture or temperature to be felt by the skin are included in the tactile feedback group.

6.3.1.1 Force Feedback

Several fundamental tasks in MOUT include force feedback as one of their central aspects. For example, firing a rifle could be taught with visual and auditory feedback only, but individuals firing actual rifles in the field must deal also with the "kick" of the weapon, the force

back into the shoulder. Thus, training without this aspect would be incomplete and unrealistic, and may not result in successful performance when rifle-firing was required in combat. It is also difficult to imagine a set of optical illusions that would allow individuals to acquire the skills for rappeling or climbing up a wall, without representing a physical rope or wall for the individual to apply the force of his weight against. These tasks can be subclassified further.

- Tasks that require the individual to pick up and manipulate an object, such as a shovel, mine, or grenade.
- Tasks in which an object must support the individual's weight.
- Tasks in which a sudden force such as weapon kick, grenade concussion, or enemy gunshot, is inflicted on the participant.

Since tasks in the first two classes of force-feedback simulation could be effected using workarounds, that is, physical objects could be integrated into the virtual world. Participants can throw solid objects with the shape and heft of grenades; for instance, and they can climb real ropes and sawhorse walls while observing the graphical presentation of a virtual environment.

The third class, however, requires powerful but controlled explosive force. In fact, simulation of concussive or "sudden force" feedback can be dichotomized into (1) those impacts that are intended to simulate somatic trauma, and (2) those that merely represent effects of an explosive force of which the participant is not the recipient. Where simulation of somatic trauma is desired, a technique that was developed for inflicting punishment by shock has been found to result in an experience resembling that of gunshot. Called the self-injurious behavior inhibiting system (SIBIS), this method has been successfully used to prevent autistic children from hurting themselves; the effect, however, is said to resemble the trauma of a shot. On the other hand, a method for simulation of grenade concussion and rifle kick has not yet been developed.

6.3.1.2 Tactile Feedback

The two basic tactile sensations are texture and temperature. VE technology has not yet progressed very far toward simulating tactile texture, but temperature could be introduced inexpensively by introduction of specialized equipment, for instance, heating elements in gloves and body suits. The tactile perception of texture, however, is certainly more important in creating the experience of immersion in the virtual world, and communicates more information about the environment to the participant.

6.3.2 Movement of the Participant Through Space

Most of the tasks included for consideration in this report require the participant to move the body through space; for example, by walking, running, crawling, jumping, rappeling, or being lifted by teammates. In VE training, this is difficult for several reasons.

- Participants are attached by wires to the apparatus.
- Space for training activities is limited.

- Position sensors cover a limited region.
- Equipment such as bulky headgear prohibits some movements.

If a task requires a marine to move across a street, for instance, which appears rather frequently in the current list, the HMD would have to be attached to a boom or other mechanical device to follow the individual (or wireless apparatus invented), the training area would need to be as wide as a street, and no currently available position sensors would be able to track movement.

Some investigators have solved these problems by devising treadmill-like platforms; these allow the participant to simulate walking and running, but are severely limited in responsiveness to the participants changes of direction and velocity. Other paradigms ask the user to run in place, or to “fly” through use of a mouse or other input device.

Other situations requiring movement of the participant through space, such as jumping from rooftops, will probably not be simulated in the near future. Again, reality could be augmented with props such as bilevel rooms, but cumbersome equipment would probably interfere with rapid or violent movements such as jumping, climbing, and fighting. Further, one of the advantages of VR for training is the portability of software, and the necessity of constructing large platforms would seriously reduce the desirability of this training method. Additional investment in research regarding the important issue of corporeal movement through space is recommended.

6.3.3 Multiple Participants

MOUT tasks generally involve multiple individuals, including team members, enemies, and bystanders. Thus, VE training will often require interaction of multiple participants. The integration of several users into a single VE, where each can see the other and all can see the environment from their own perspectives, and where consequences of one's behavior, such as drawing sniper fire, affect the others, offers great potential for training. On the other hand, multi-user VEs present technological challenges in terms of software, hardware, and computational load. At this stage of development, researchers have barely begun to consider solutions to the problems of integrating multiple participants. The MOUT program presents opportunities for pursuing these challenges, which must be met if team activities are to be simulated.

6.3.4 Simulating Human Behavior

MOUT tasks include situations involving cooperation with team members, adversarial interaction with enemies, and assessment of and interaction with noncombatants. Thus, it is important to represent simulated human behavior in the virtual world. Research by RTI staff has determined that virtual environments might be arrayed on a continuum of interactivity or agency of virtual actors.

- Environments with no people: Most present VE's are of this type, which is limited in terms of training applications and interest to the participant.

- VE's with static, nonmoving people: Properly implemented, these can populate a battlefield or urban environment when used appropriately with more lifelike virtual persons.
- Preprogrammed moving people with absolute spatial coordinates: These suggest use of the computer as a movie projector. Again, these could enhance a populated scene with little added computing cost.
- Preprogrammed moving people with spatial coordinates relative to the participant: These appear in a script regardless of the participant's location and orientation, without interacting with the trainee.
- Task-oriented reactive virtual persons (VPs). Even simple task cooperation requires sophisticated artificial intelligence (AI) to encode the participant's movements into meaningful units and select a reaction.
- Socioemotional reactive VP's: Allows the introduction of normative social influence in the virtual context.
- Interreactive VP's: movements coordinated among multiple virtual persons: These sophisticated cellular automata might populate a scene, spur one another on as adversaries or comrades, create random disturbances and distractions, and otherwise mimic dynamic social behavior.
- Avatars: real people encoded in real time. Some observers (e.g., Neal Stephenson) assume that the operation of avatars will provide the greatest benefit in the use of VE methods. They are best considered to be a gateway to the programming of artificially intelligent realistic virtual actors.

Various researchers are accomplishing much toward populating virtual worlds. There are, of course, many complexities in developing algorithms for programming the representation of human flesh and movement, and while considerable progress has been made, there is still a long way to go before realistic human behavior can be modeled. This condition does not mean, however, that virtual worlds cannot be populated in a practical sense. Graphical images of humans can be programmed into VR scripts without much difficulty, using presently available technology; preprogrammed human behavior can be represented, but simulated human agency (or realistic randomized behavior) is still in the future.

Psychologists and computer scientists are very interested in simulating human behavior in virtual worlds. Knowledge based experts are familiar with algorithms for modeling thought, but the graphical simulation of human behavior is new territory. While scientists have experience with abstract simulations of cognitive and societal dynamics, the research discussed here would bring these to a much more concrete level.

For training, it is apparent that virtual actors can function as coaches/tutors, enemies, team members, and noncombatants; training paradigms in which the virtual person substitutes for the participant in vicarious performance of behaviors, allowing a third-person analysis of skills can

be visualized. While realistic persons might enhance the scene, suprealistic persons can be used effectively for training. These individuals can respond to the participant's behavior in exaggerated ways, indicating efforts and correct behaviors in ways that real people could not do. In summary, the production of virtual actors will become a very important element of training in virtual worlds.

6.3.7 Explosions

In MOUT combat, the individual is confronted with mines, grenades, C-4, TNT, and other explosive devices and materials. The training task in some cases comprises learning to emplace the explosives, and in other cases learning to avoid being injured by the blast. To simulate these situations, a good explosion algorithm must be devised, taking into account a number of parameters, so that the effects of the explosion are similar to the effects of real ones.

Parameters that affect the result of an explosion include not only innate factors such as type, amount, and placement of explosive, but also hardness of walls and other reflecting obstacles, contents of rooms and their hardness, angles of reflection, and trajectory arc as a function of gravity and mass of debris. Thus, programming of a realistic explosion algorithm will not be trivial; but it is necessary to ensure that individuals are trained to expect realistic consequences from explosives.

6.3.8 Field of View

Some tasks allow the individual to look directly at the object of interest, while others require attending to peripheral events as well. This consideration will affect the selection of equipment, notably it will force a choice between HMDs with fine resolution and narrow field of view, and those with lower resolution and larger field of view.

6.3.9 Perceptual Versus Motor Emphasis

Some MOUT tasks are mainly perceptual: the individual's task is to reconnoiter, or to identify booby traps, or plan a route to a strategic position. Others are mainly motor tasks, such as climbing walls, throwing grenades, and fortifying positions. Of course, many of the tasks include both perceptual and motoric elements.

Considering the input/output exchange between the participant and the VE computer, perceptual tasks are considered to be ones that emphasize computer output, and motor ones as tasks emphasizing computer input from the participant. This view indicates the prescription for the allocation of computer resources in the programming of the virtual environment.

6.3.10 Summary of MOUT Research Issues

In summary, a number of issues can be addressed within a comprehensive MOUT research program. Hardware issues include:

- Development of tactile and force-feedback apparatus.
- Increasing the peripheral view of virtual environment headgear.
- Increasing the polygonal throughput of image generators.

- Integration of workarounds (actual physical objects) into a virtual reality paradigm.
- Infliction of painful but nonthreatening consequences (i.e., simulated shrapnel and gunshot) for errors in precarious situations.
- Creation of weapons that are cut away to allow fitting to HMD, within which virtual sights would be presented.
- Development of apparatus to allow simulation of corporeal movement (i.e., walking, running, crawling, etc.) within a constrained area.
- Force-feedback methods to simulate the kick of a weapon and other concussions.
- Development of wireless, lightweight apparatus to allow flexibility of movement.

Software issues include:

- Representing workarounds, the location of physical objects in the workspace, in the virtual environment.
- Algorithms for simulating agentic or randomized human behavior in virtual actors.
- Integrating multiple participants interactively into a VE simulation.
- Scripts: programmed sequences of events must be logically consistent, but present novelty and surprise.
- Development of algorithms for modeling explosions.
- Algorithms to give the illusion of movement (i.e., climbing a rope, falling, etc.) with reduced actual corporeal motion.

Training issues include:

- Whether degrees of photographic fidelity of representations affect transfer of training from virtual to actual tasks.
- Whether stress can be induced in VE simulations, and secondly whether stress during training enhances or diminishes training effectiveness.
- How the participant's sense of "immersion," or acceptance of the VE illusion, affects training outcomes.
- How training is affected by the use of body motion to accomplish motor tasks during an exercise, as opposed to sitting in one place working a joystick or other control.

- Whether interactivity, or realistic dynamics between participant and environmental objects, affects training.
- Interactions among these factors.

References

- Aukstakalnis, S., & Blatner, D. (1992). *Silicon mirage: The art and science of virtual reality*. Berkeley: Peachpit Press, Inc.
- Badler, B., Phillips, C., & Weber, B. (1993). *Simulating humans: Computer graphics animation and control*. New York: Oxford University Press.
- Biocca, F. (1992). Will simulation sickness slow down the diffusion of virtual environment technology? *Presence: Teleoperators and Virtual Environments*, 1(3).
- Brooks, F., Airey, J., Alspaugh, J., Bell, A., Brown, R., Hill, C., Nimscheck, U., Rheingans, P., Rohlf, J., Smith, D., Turner, D., Varshney, A., Wang, Y., Wever, H., & Yuan, X. (1992). Six generations of building walkthroughs (TR-92-026). Chapel Hill, NC: University of North Carolina, Department of Computer Science.
- Calvin, J., Dickens, A., Gaines, B., Metzger, P., Miller, D., & Owen, D. (1993). The SIMNET virtual world architecture. *1993 IEEE Virtual Reality Annual International Symposium*. IEEE Press.
- Chen, D., Pieper, S., Singh, S., Rosen, J., & Zeltzer, D. (1993). The virtual sailor: An implementation of interactive human body modeling. *1993 IEEE Virtual Reality International Annual Symposium*. IEEE Press.
- Cholewiak, R., & Collins, A. (1991). Sensory and physiological bases to touch. In Heller, M. & Schiff, W. (Eds.). *The psychology of touch*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Deering, M. (1993). Data complexity for virtual reality: Where do all the triangles go? *1993 IEEE Virtual Reality Annual International Symposium*. IEEE Press.
- Defanti, T., Sandim, D., & Cruz-Neira, C. (1993). A "room" with a "view." *IEEE Spectrum*, 30(10).
- Eberhart, R., & Kizakevich, P. (1993). Determining physiological effects of using VR equipment. *Proceedings of the Virtual Reality for Disabled Persons Conference, San Francisco, CA*.
- Geppert, L. (1993). Not your father's CPU. *IEEE Spectrum*, 30 (12).
- Gerber, S. (1974). *Introductory hearing science*. Philadelphia: W.B. Saunders Company.
- Hardenbergh, H. (1994). CPU performance: Where are we headed? *Dr. Dobbs Journal*, 19 (1).
- Hayward, T. (1993). *Adventures in virtual reality*. Carmel, IN: Que Publishing.
- Isdale, J. (1993). Cyberspace development kit. *Software Development*, 1(6).

- Kizakevich, P., Jochem, W., & Beadles, R. (1989). Ambulatory monitoring of cardiac and respiratory activity. *Proceedings of the 1989 EPAIA & WAM Symposium on Total Exposure Assessment Methodology*, Las Vegas, Nevada.
- Kizakevich, P., McCartney, M., Jochem, W., Raymer, J., & Pellizzari, E. (1993). Estimating minute ventilation from activity-related cardiac measurements. *Journal of Clinical Engineering*. 18(3).
- Latham, R. (Ed.) (1993a). Head mounted display survey. *Real Time Graphics*. 2(2).
- Latham, R.(Ed.) (1993b). Kaiser's VIM head mounted display provides mosaicked LCDS. *Real Time Graphics*. 2(3).
- Latham, R. (Ed.) (1993c). Survey of high performance image generators. *Real Time Graphics*. 2(5).
- Latham, R. (Ed.) (1993d). Progress in Standards for Networked Simulation: A Report on the Ninth DIS Workshop. *Real Time Graphics*. 2(5).
- Lawrence, A., De Rossi, D., & Baughman, R. (1993). Applications of conductive polymers in medical robotics and prosthetics. In *Proceedings of the 15th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE Press.
- Lee, K. (1989). *Automatic Speech Recognition*. Boston: Kluwer Academic Publishers.
- McCauley, M., & Sharkey, T. (1992). Cybersickness: Perception of self-motion in virtual environments. *Presence: Teleoperators and Virtual Environments*. 1(3).
- McKenna, M., & Zeltzer, D. (1992). Three dimensional display systems for Virtual environments. *Presence: Teleoperators and Virtual Environments*. 1(4).
- Meyer, K., Applewhite, H., & Biocca, F. (1992). A survey of position trackers. *Presence: Teleoperators and Virtual Environments*. 1(2).
- Monkman, G. (1992). An electrorheological tactile display. *Presence: Teleoperators and Virtual Environments*, 1(2).
- Newman, A. (1985). Development of medical electronic devices in the APL space department. *Johns Hopkins APL Technical Digest*, 6(1).
- Slater, M., & Usoh, M. (1993). Presence in immersive virtual environments. *1993 IEEE Virtual Reality Annual International Symposium*. IEEE Press.

Small, C. (1993). Workstations vs. PCs. *EDN*, 38(6).

Standish, LCD. (1991). Liquid crystal display products. *Product Brochure 6/91-10M*.

Watt, A., & Watt, M. (1992). *Advanced animation and rendering techniques: Theory and practice*. New York: ACM Press.

Wenzel, E. (1992). Localization in virtual acoustic displays. *Presence: Teleoperators and Virtual Environments*, 1(1).

Werner, K. (1993). The flat panel's future. *IEEE Spectrum*, 30(11).

Zeltzer, D., & Johnson, M. (In press). Virtual actors and virtual environments: Defining, modeling, and reasoning about motor skills. In MacDonald, L., & Vince J. (Eds.). *Interacting with virtual environments*. John Wiley and Sons.

Appendix A

Analysis of Training Requirements for Movement Through Urban Areas Outside of Buildings

Analysis of Training Requirements for Movement Through Urban Areas Outside of Buildings

A.1 Fundamental skills

A.1.1 Cross a wall

Sensory: Visual, kinesthetic.

Perceptual: Individuals must be able to see the wall and use other sensory cues to provide information about their movements and positions relative to the wall.

Cognitive: This is a motor task.

Motor: Individuals must use arms and legs to pull themselves over the wall.

Instructional/Training: This should be practiced in exercises.

A.1.2 Look and move around a corner

Sensory: Vision and hearing inform the individual about consequences of moving around the corner.

Perceptual: Individuals must perceive their position relative to potential enemies and cover.

Cognitive: The individual must be able to imagine where adversaries could be hiding.

Motor: This is not, in itself, a specialized task.

Instructional/Training: This simple skill does not require special training.

A.1.3 Move past a ground floor window

Sensory: See the window and where the individual wants to go.

Perceptual: The individual must recognize the potential for danger from within the window.

Cognitive: The individual must judge whether it is important to pass the window, given the degree of risk involved.

Motor: It is necessary to duck or crawl when passing a ground-floor window, which might contain adversaries.

Instructional/Training: Trainees should pass windows of different heights, widths, etc., with consequences; that is, a role-playing adversary should give them feedback when they are seen from within.

A.1.4 Exit a doorway

Sensory: Vision and hearing relay information about what dangers lie outside the doorway.

Perceptual: The individual must look out.

Cognitive: Vigilance, attention to cues suggesting the presence of danger.

Motor Response: In itself, this is not a complicated motion.

Instructional/Training: This skill should be taught in a context.

A.1.5 Move parallel to a building

Sensory: Visions and hearing input information about adversaries and the environment.

Perceptual: Individuals should be able to see a semicircular area around themselves.

Cognitive: N/A

Motor Response: Again, this is not a specialized behavior.

Instructional/Training: This should be trained in the context of performing a larger task.

A.1.6 Cross an open area

Sensory: Must rotate head and body to view 360 degrees.

Perceptual: The individual looks around for signs of danger and to assess position relative to goal.

Cognitive: Judgment is required in determining risks involved in crossing the open area.

Motor Responses: This is, in itself, not a unique behavior.

A.1.7 Move across rooftops

Sensory: This task requires vision and good kinesthetic sense for balancing.

Perceptual: Good depth perception is required for jumping from one roof to another.

Cognitive: N/A

Motor: This is a perceptual/cognitive task.

Instructional/Training: This skill is probably not trained per se, but rather will develop as individuals learn how to use cover, and what their own field of fire is from various positions.

A.1.8 Select and take up hasty firing positions

Sensory: Individuals must see clearly around themselves.

Perceptual: Must see good positions to hide in, and estimate the location of the adversary.

Cognitive: Fast response is necessary for this task: attention must be tuned to relevant details, judgments must be quick and correct, and movements must be immediate and self-assured.

Motor: This might require some contortion, jumping, bending, etc.

Instructional/Training: Individuals should prepare for this task by participating in realistic war games.

A.2 Avoid open areas

A.2.1 Locate enemy locations and estimate the enemy's fields of fire

Sensory: Visual and auditory senses are involved in this task.

Perceptual: The individual must be able to see enemy positions and hear movement, gunfire, etc.

Cognitive: Attention must be focused on the task, and perceptions tuned by knowledge of combat techniques. That is the enemy can be located better if the individual knows what kinds of positions they might select. Estimation of fields of fire requires inferential processes and the ability to imagine the perspective of the enemy in order to infer the trajectory of a missile originating at the enemy's location.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Individuals should be trained in judging what positions an enemy would select to hide in. Further, they should be trained in estimating the path that a bullet might take, from a given position, by observing and predicting.

A.2.2 Identify and contrast areas that provide cover and those that are open relative to the enemy's location and fields of fire

Sensory: Visual sense required.

Perceptual: The individual must have a clear view of the area.

Cognitive: Knowledge of types of cover is necessary for this task; it is also necessary to be able to take various hypothetical perspectives in order to estimate the enemy's fields of fire.

Motor: This is a perceptual/cognitive task.

Instructional/Training: This skill is probably not trained per se, but rather will develop as individuals learn how to use cover and what their own field of fire is from various positions.

A.2.3 Select and validate areas that provide cover for movement

Sensory: Visual sense required.

Perceptual: The individual must be able to perceive an array of areas with potential to provide cover, and must see these well enough to estimate their value for that purpose.

Cognitive: Requires knowledge of types of cover, including the factors by which areas may be ranked in terms of providing good cover. Then, the individual must compose a mental list of the available areas, rank-ordered, and select the best one(s).

Motor: This is a cognitive task.

Instructional/Training: Individuals should be given a clear set of rules delineating the virtues of areas that provide cover. This list of 5-10 items should be memorized, perhaps using a mnemonic strategy, so that it can be recalled under pressure. Further, observation of movies and participation in realistic field exercises might make the impact of the information more salient.

A.3 Conduct movement using cover

A.3.1 Through underground structures

Sensory: Requires visual, tactile, and auditory sensation.

Perception: Unless the terrain is very regular and/or familiar, lighting and other perceptual aids (tapping sticks, infrared visors, etc.) will be necessary.

Cognitive: This task requires several unusual cognitive skills. The unavailability of mapped landmarks will require individuals to maintain knowledge of their location, using compass etc.; it is imperative that individuals not become disoriented in the dark or claustrophobic. Attention must be focused on those senses (hearing, touch), which work well in the darkness, and all input sense data must be analyzed for cues to presence of enemies and proximity to goal.

Motor: Requires walking over uneven underground terrain in the dark without falling and stumbling.

Instructional/Training: Individuals should be trained to walk over uneven ground in darkness, with unexpected "ambushes" focusing their attention on the danger of the situation.

A.3.2 Through or behind secured buildings

Sensory: Visual/auditory senses required.

Perceptual: The individual must be able to detect the presence of danger from without the secured area.

Cognitive: Must have knowledge of the extent of the secured zone, and possible sources of attack beyond it.

Motor: This resembles ordinary movement, except that the individual will want to keep out of sight or line of fire of enemies positioned around the perimeter of the secured area.

Instructional/Training: This skill should be acquired through training in more dangerous maneuvers. Good habits built up in more stringent exercises should transfer to this.

A.3.3 Through streets using cover that may be available such as rubble, cars, trees, etc.

Sensory: Visual sense required.

Perceptual: Individuals must be able to see the terrain through which they are moving, including sources of cover as well as obstacles on the ground that may trip or impede them.

Cognitive: Must be able to judge the enemy's position and possible areas of vulnerability to fire. The individual must also be able to evaluate the various sources of cover.

Motor: Requires running, possible carrying a load, ducking, climbing, crawling, etc.

Instructional/Training: This should be trained in realistic interaction with mock enemies who will shoot at any target that reveals itself.

A.3.4 Integrate movement with vehicles and tanks

Sensory: Visual, auditory senses used.

Perceptual: The individual must perceive the moving vehicle as well as any features of the terrain that may result in a change of velocity.

Cognitive: Individuals must be able to visualize memorized terrain on the opposite side of the vehicle or tank, in order to place themselves in a position relative the vehicle that shield them from fire. Further, some mental calculus is required in order to time personal movements with those of vehicles.

Motor: This requires running, ducking, and coordination of one's movements to those of another object.

Instructional/Training: This can be trained on a realistic training ground with mock enemies firing at any revealed target.

A.3.5 Vary speed of movement

A.3.6 Maintain dispersion

A.4 Suppress or obscure enemy fires

A.4.1 Enemy fires brought onto friendly soldiers

A.4.2 Friendly soldiers identify enemy fires, to include location

A.4.3 Friendly soldiers bring fires or means to obscure them such as smoke to bear on enemy fires to suppress or obscure them

A.4.4 Friendly soldiers are able to determine that enemy fires have been suppressed or obscured

A.5 Move at night or during periods of reduced visibility

A.5.1 Night conditions under varying levels of light

Sensory: Visual, auditory, and tactile senses required.

Perceptual: The challenges of this task are perceptual: moving in the dark requires some way to gain information about the terrain. This information may be acquired through use of high-tech visual aids (infrared helmets) or through carefully tapping along with a stick or other probe. Note that one challenge is simply to determine a secure place to set one's foot for the next step, but a more serious challenge is to determine in the dark where an enemy might be hiding.

Cognitive: Individuals walking in the dark must be attentive to possible dangers, both of attack and of immediate obstacles. Further, skill in map-following and identifying directions from the stars and lighted landmarks is useful.

Motor: Walking, crawling, etc., as usual, but movement over unseen ground, through obstacles that are not clearly visible.

Instructional/Training: If special visual equipment is to be used, its implementation should be learned through hands-on practice. Otherwise, training on obstacle courses at night should train this skill, or at least impart some familiarity with it.

A-5.2 Reduced visibility such as fog

Sensory: Auditory and tactile senses can be used.

Perceptual: Fog (and smoke, etc.) reduces the perceptual options considerably. This situation differs from nighttime travel in that infrared instruments will not penetrate it.

Cognitive: The individual must maintain a sense of location, despite inability to identify any landmarks.

Motor: Walking, running, crawling over unknown terrain (assuming that the individual can not clearly see the ground).

Instructional/Training: Unless fog-penetrating technologies are to be employed, this task could be trained by blindfolded navigation of an obstacle course.

A.6 Select routes that will not mask friendly fires

A.6.1 Location of friendly weapon systems

Sensory: Visual and auditory senses.

Perceptual: The individual must be able to see where weapon systems are installed (if that is the mode of information) or be able to hear the communication of the information clearly (walkie-talkie, briefing, etc.).

Cognitive: If friendly weapon systems are hidden, the individual must be able to remember where they are, based upon briefings and other communications. This requires more than memory of verbal descriptions: it requires translation of verbal communications into geographical coordinates and positions.

Motor: This is primarily a cognitive task.

Instructional/Training: Individuals must be schooled, with rote memory, in the meanings of terms that might be used to convey the positions of friendly weapon systems. Further, when visual cues are used, the individual must be trained in identification of features of weapons that distinguish friendly ones from those possessed by the enemy.

A.6.2 Trajectories and impact points of projectiles being fired from friendly weapon systems

Sensory: If these are small arms, visual and auditory cues are used.

Perceptual: When line-of-sight firearms are used, the individual must see the weapon, including the direction it is aimed. Further, auditory perception is an important cue that a weapon is being fired.

Cognitive: Estimation of trajectories and impact points is often no more than simple straight-line extrapolation. In some cases, however (i.e., aircraft and mortar fire), estimation of impact points requires curvilinear estimation, and inside knowledge of the aim and target of the friendly fire.

Motor: This is a perceptual/cognitive task.

Instructional/Training: When individuals learn to employ weapons themselves, they will acquire knowledge of estimation of missile trajectories.

A.6.3 Ability to determine if routes will mask friendly fires

Sensory: Visual, auditory senses used.

Perceptual: When line-of-sight firearms are used, the individual must see friendly weapons, including the direction they are aimed. Further, auditory perception is an important cue that a weapon is being fired.

Cognitive: Trajectories and impact points must be estimated. In this case, however, these must be estimated in relation to movement of the individual.

Motor: This is a perceptual/cognitive task.

Instructional/Training: When individuals learn to employ weapons themselves, they will acquire knowledge of estimation of missile trajectories and how to avoid obfuscation of these.

A-7 Cross open areas such as streets, fields, open areas between buildings, etc., rapidly under concealment of fires, suppressive fires, and/or smoke

A.7.1 Identify open areas

Sensory: Primarily visual.

Perceptual: The individual must visually perceive the layout of the environment, including a path to cross the area and possible enemy positions.

Cognitive: The individual must estimate the amount of vulnerability versus cover in the open area, the amount of time required to cross it, and the probability of risk.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Hiking and war games should provide training for this skill.

A.7.2 Contrast open areas with enemy positions

Sensory: Primarily visual.

Perceptual: The individual must visually perceive the layout of the environment, including a path to cross the area and possible enemy positions.

Cognitive: The individual must estimate the amount of vulnerability versus cover in the open area, the amount of time required to cross it, and the probability of risk.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Hiking and war games should provide training for this skill.

A.7.3 Select routes that are concealed

Sensory: Primarily visual.

Perceptual: The individual must be able to visually perceive the terrain between the beginning and goal positions.

Cognitive: Where the entire route cannot be seen, the individual must be able to interpolate probable concealment versus vulnerability.

Motor: This is a perceptual/cognitive task.

Instructional/Training: This should be trained through interaction with mock enemies who will fire at any revealed target.

A.7.4 Plan and place fires on areas that cannot be concealed

A.7.5 Call for or Time the Execution of These Fires to Support the Crossings of These Areas by Friendly Soldiers

Sensory: Visual sensation required.

Perceptual: The individual must be able to see the clear area, including the movements of those crossing it.

Cognitive: Individuals must be able to estimate the movements of others through the open area, as well as the trajectories of their own fires.

Motor: Firing weapons requires some motor skills.

Instructional/Training: This could be trained in practice in realistic or virtual (interactive) situations.

A.8 Move on roof tops that are not covered by direct enemy fires

A.8.1 Identify roof tops that potentially are not covered by direct enemy fires

Sensory: Visual sense is used.

Perceptual: The individual must be able to see the roof top and its relation to enemy positions.

Cognitive: Must infer probable locations of enemies and their line of fire, to judge whether they can hit the roof top or not.

Motor: This is a perceptual/cognitive task.

Instructional/Training: This should be trained by practice with interactive mock enemies.

A.8.2 Contrast these roof tops with enemy locations to determine the ability of the enemy to place fires on them

Sensory: Primarily visual sense required.

Perceptual: The individual must be able to see the roof top and its relation to enemy positions.

Cognitive: Must infer probable locations of enemies and their line of fire, to judge whether they can hit the roof top or not.

Motor: This is a perceptual/cognitive task.

Instructional/Training: This should be trained by practice with interactive mock enemies.

A.8.3 Select routes that are concealed from enemy direct fires

Sensory: Primarily visual.

Perceptual: The individual must be able to see the route and its relation to enemy positions.

Cognitive: Must infer probable locations of enemies and their line of fire, to judge whether they can hit the route or not.

Motor: This is a perceptual/cognitive task.

Instructional/Training: This should be trained by practice with interactive mock enemies.

A.8.4 Move over the selected routes

Sensory: Visual, tactile, kinesthetic, and proprioceptive senses involved.

Perceptual: The individual must be able to see where he is going, including any weaknesses or flaws in the supporting structure.

Cognitive: Must be able to judge height from the ground and positions of enemies relative to self

Motor: Walking, running, jumping, crawling, balancing.

A.9 Select subsequent positions before moving

A.9.1 Plan and observe routes over which movement is to take place

Sensory: Primarily visual.

Perceptual: First, must be able to see and hear plans, such as maps (including those drawn in the dust) and instructions. Second, must be able to visually observe the route over which movement is to take place.

Cognitive: Knowledge of terrain, and knowledge, inference, or estimation of enemy positions and their line of fire are necessary.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Individuals should be trained in the ranking of factors that favor one route over another. Further, field practice with mock enemies will instill a sense of how to avoid danger.

A.9.2 Contrast routes and planned positions with enemy locations and fields of observation and fires

Sensory: Primarily visual.

Perceptual: First, must be able to see and hear plans, such as maps (including those drawn in the dust) and instructions. Second, must be able to visually observe the route over which movement is to take place and to observe probable locations of enemies.

Cognitive: Knowledge of terrain, and knowledge, inference, or estimation of enemy positions and their line of fire are necessary, as is judgment of risks involved with each route.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Individuals should be trained in the ranking of factors that favor one route over another. Further, field practice with mock enemies will instill a sense of how to avoid danger.

A.10 Move around corner of building

A.10.1 Soldier lays flat on ground

Sensory: Kinesthetic, tactile, and other senses are involved in this.

Perceptual: The individual's point of view is strictly constrained by proximity to the ground.

Cognitive: This is a motor task that has been selected by previous cognitive processes.

Motor: This task requires motor coordination and the ability to lie down from a standing position.

Instructional/Training: This task should be learned in obstacle courses and mock combat.

A.10.2 Soldier slowly crawls as close to the corner of the building as possible without being exposed

Sensory: Kinesthetic and visual senses are featured.

Perceptual: The individual sees the corner of the building and sees and feels the ground.

Cognitive: Same as previous.

Motor: Crawling, dragging the body along the ground.

Instructional/Training: This task should be learned in obstacle courses and mock combat.

A.10.3 While maintaining the weapon to the front, the soldier flares the elbows and pushes the body forward with the toes to peep around the building

Sensory: Tactile and visual senses predominate.

Perceptual: The individual feels the ground and is aware of body movements and visually tries to see around the corner.

Cognitive: Same as previous.

Motor: This motor task requires crawling and stretching out the neck.

Instructional/Training: This exercise should be practiced in mock battle situations, where “enemies” will fire on any target that is seen.

A.11 Moving past first story windows

A.11.1 Identify windows

Sensory: Visual sense required.

Perceptual: Must be able to see the side of the building.

Cognitive: The individual must have knowledge of where windows are located.

Motor: This is a perceptual task.

Instructional/Training: Individuals will probably not have to be trained to identify windows.

A.11.2 Move along the wall to the window while avoiding rubbing equipment against the wall which might alert the enemy

Sensory: Tactile, kinesthetic, visual, and auditory senses are involved.

Perceptual: The individual must be able to see and feel the wall and judge position of body and equipment relative to it.

Cognitive: Must judge what features of the topography will result in equipment noise.

Motor: This task requires motor coordination, walking near the wall without banging against it.

Instructional/Training: This maneuver should be practiced with full equipment, perhaps in mock combat where mistakes have immediate consequences.

A.11.3 Bend down and move under the window

Sensory: Visual and kinesthetic senses are involved.

Perceptual: Individuals must see the position of the window and visualize the position of their bodies relative to it, and must perceive their own center of gravity, with equipment, in order to maintain balance.

Cognitive: This is primarily a motor task.

Motor: Walking while bent over with heavy equipment.

Instructional/Training: This movement should be practiced.

A.12 Moving past basement windows

A.12.1 Identify windows

Sensory: Visual sense required.

Perceptual: Must be able to see the side of the building.

Cognitive: The individual must have knowledge of where windows are located.

Motor: This is a perceptual task.

Instructional/Training: Individuals will probably not have to be trained to identify windows.

A.12.2 Move along the wall to the window while avoiding rubbing equipment against the wall which might alert the enemy

Sensory: Tactile, kinesthetic, visual, and auditory senses are involved.

Perceptual: The individual must be able to see and feel the wall and the position of body and equipment relative to it.

Cognitive: Must judge what features of the topography will result in equipment noise.

Motor: This task requires motor coordination, walking near the wall without banging against it.

Instructional/Training: This maneuver should be practiced with full equipment, perhaps in mock combat where mistakes have immediate consequences.

A.12.3 Step or jump over the window; avoid exposing the legs

Sensory: Visual and kinesthetic senses are used in this task.

Perceptual: Individuals must see the width of the window and judge their own weight and balance for jumping.

Cognitive: The individual must be able to estimate the perspective of enemies inside the window.

Motor: This is a motor task and requires stepping or jumping with heavy equipment.

Instructional/Training: This task should be practiced with equipment, in mock combat situation.

A.13 Crossing a fence or wall

A.13.1 Make a visual reconnaissance of the fence or wall to identify booby traps, the next position to move to, and to look for enemy activity

Sensory: This task uses visual and auditory modalities.

Perceptual: The individual must be able to see the fence or wall, listen for the presence of enemies, and be able to see the next position to move to, as well as any enemy positions that might overlook it and the path to it.

Cognitive: The person must have knowledge of booby trap types and of virtues of various types of positions.

Motor: This is a perceptual/cognitive task.

Instructional/Training: The individual should be trained in booby trap methods and in techniques for moving from one place to the other.

A.13.2 Quickly roll over the top of the wall; keep the body silhouette as low as possible

Sensory: This task requires visual and kinesthetic sensation.

Perceptual: Must see the wall and what is on the other side of it, as well as feeling body position in order to retain control of movements.

Cognitive: Must be able to estimate the extent of silhouette visibility.

Motor: This is primarily a motor task.

Instructional/Training: Individuals should practice rolling over walls and fences of various kinds.

A.13.3 Penalize the soldier for not locating booby traps, enemy locations or for exposing the body silhouette to enemy observation

Sensation: Same as previous.

Perception: Same as previous.

Cognition: This task is defined as a learning exercise: the individual learns to associate consequences with mistakes.

Motor: Same as previous.

Instructional/Training: This is a training suggestion. Once individuals have learned to roll over walls, they should practice this in mock combat situation, with real or virtual enemies who will shoot them (paint, lasers, etc.) when they expose themselves or fail to locate booby traps.

A.14 Passing or exiting doorways

A.14.1 Make a visual reconnaissance of doorway to identify booby traps, the next position to move to, and to look for enemy activity

Sensory: This task uses visual and auditory modalities.

Perceptual: The individual must be able to see the fence or wall, listen for the presence of enemies, and be able to see the next position to move to, as well as any enemy positions that might overlook it and the path to it.

Cognitive: The person must have knowledge of booby trap types and of virtues of various types of positions.

Motor: This is a perceptual/cognitive task.

Instructional/Training: The individual should be trained in booby trap methods and in techniques for moving from one place to the other.

A-14.2 Quickly move through the doorway to the next position; keep the body silhouette as low as possible

Sensory: This task requires visual and kinesthetic sensation.

Perceptual: Must see the doorway and what is on the other side of it, as well as feeling

Cognitive: Must be able to estimate the extent of silhouette visibility and to select a next position.

Motor: This is primarily a motor task.

Instructional/Training: Individuals should practice rolling over walls and fences of various kinds.

A.14.3 Penalize the soldier for not locating booby traps, enemy locations or for exposing the body silhouette to enemy observation

Sensation: Same as previous.

Perception: Same as previous.

Cognitive: This task is defined as a learning exercise: the individual learns to associate consequences with mistakes.

Motor: Same as previous.

Instructional/Training: This is a training suggestion. Once individuals have learned to roll over walls, they should practice this in mock combat situation, with real or virtual enemies who will shoot them (paint, lasers, etc.) when they expose themselves or fail to locate booby traps.

A.15 Movement in streets

A.15.1 When possible, conduct movements inside buildings

Sensory: All.

Perceptual: Includes all forms of perception.

Cognitive: This is simply a bit of information that the individual should have available for recall.

Motor: All.

Instructional/Training: This opinion should be conveyed, memorized, and tested.

A.15.2 Make a visual reconnaissance to identify positions that provide maximum cover and concealment

Sensory: Primarily visual.

Perceptual: This is primarily a perceptual task: the individual looks around the area to identify a route to take.

Cognitive: Requires knowledge of qualities of cover and concealment, as well as some knowledge of likely enemy positions.

Motor: This is a perceptual task.

Instructional/Training: Should be trained with other tasks employing cover and concealment.

A.15.3 Select route that provides cover and concealment

Sensory: Visual sense used primarily.

Perceptual: The individual must be able to see possible cover and should try to see likely enemy positions along the various routes.

Cognitive: Requires knowledge of cover and concealment, as well as knowledge to like hiding places of the enemy.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Individuals should learn how to judge the dimensions of cover, through classroom and literature, and further through practice in mock combat.

A.15.4 Use smoke and suppressive fires if a route with cover and concealment is not available

Sensory: Visual.

Perceptual: The person must visually select a route to take.

Cognitive: Must be able to judge when cover and concealment are not available. This task also requires knowledge of techniques for creating smoke and suppressive fires.

Motor: The individual must fire a weapon or create smoke.

Instructional/Training: Individuals should be trained by practice with smoke bombs, and with use of weapons.

A.15.5 Penalize the soldier for violating these principles

Sensory: Same as previous.

Perceptual: Same as previous.

Cognitive: This task is defined as a learning exercise: the individual learns to associate consequences with mistakes.

Motor: Same as previous.

Instructional/Training: This is a training suggestion. Once individuals have learned to move in streets, they should practice this in mock combat situation with real or virtual enemies who will shoot them (paint, lasers, etc.) when they expose themselves.

A.16 Movement across open areas

A.16.1 Make a visual reconnaissance to identify positions that provide maximum cover and concealment

Sensory: Primarily visual.

Perceptual: This is primarily a perceptual task: the individual looks around the area to identify a route to take.

Cognitive: Requires knowledge of qualities of cover and concealment, as well as some knowledge of likely enemy positions.

Motor: This is a perceptual task.

Instructional/Training: Should be trained with other tasks employing cover and concealment.

A.16.2 Select route that provides cover and concealment

Sensory: Visual sense used primarily.

Perceptual: The individual must be able to see possible cover and should try to see likely enemy positions along the various routes.

Cognitive: Requires knowledge of cover and concealment, as well as knowledge to like hiding places of the enemy.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Individuals should learn how to judge the dimensions of cover, through classroom and literature, and further through practice in mock combat.

A.16.3 Use smoke and suppressive fires if a route with cover and concealment is not available

Sensory: Visual.

Perceptual: The person must visually select a route to take.

Cognitive: Must be able to judge when cover and concealment are not available. This task also requires knowledge of techniques for creating smoke and suppressive fires.

Motor: The individual must fire a weapon or create smoke.

Instructional/Training: Individuals should be trained by practice with smoke bombs and with use of weapons.

A.16.4 Move across open areas rapidly from position to position without masking suppressive fires

Sensory: Primarily visual.

Perceptual: Must be able to see potential enemy positions, as well as goal.

Cognitive: This task requires suppression of fear.

Motor: Running, carrying equipment, looking around.

Instructional/Training: This task should be trained in mock combat with consequences for failure to observe important cues.

A.16.5 When the position is reached, be prepared to cover the movement of the other members of the team

Sensory: Visual, kinesthetic.

Perceptual: The individual must be able to see the movement of team members and possible enemy positions.

Cognitive: Must be able to infer locations of enemies.

Motor: Turning head, holding and firing weapon.

Instructional/Training: This should be practiced in mock combat, with consequences for failure to perform adequately.

A.16.6 Penalize the soldier for violating these principles

Sensory: Same as previous.

Perceptual: Same as previous.

Cognitive: This task is defined as a learning exercise: the individual learns to associate consequences with mistakes.

Motor: Same as previous.

Instructional/Training: This is a training suggestion. Once individuals have learned to move in streets, they should practice this in mock combat situation, with real or virtual enemies who will shoot them (paint, lasers, etc.) when they expose themselves.

A.17 Select, occupy, and use a hasty firing position during movement

A.17.1 Corners of buildings

Sensory: Visual sense used primarily.

Perceptual: Must be able to see the position.

Cognitive: This task requires quick reactions as opposed to thinking.

Motor: Turning body vigilantly, holding and firing weapon.

Instructional/Training: This task should be trained by practice in mock combat situation with consequences for mistakes.

A.17.2 Behind a wall

Sensory: Visual sense used primarily.

Perceptual: Must be able to see the position.

Cognitive: This task requires quick reactions as opposed to thinking.

Motor: Turning body vigilantly, holding and firing weapon.

Instructional/Training: This task should be trained by practice in mock combat situation with consequences for mistakes.

A.17.3 Rubble

Sensory: Visual sense used primarily.

Perceptual: Must be able to see the position.

Cognitive: This task requires quick reactions as opposed to thinking.

Motor: Turning body vigilantly, holding and firing weapon; walking over uneven ground.

Instructional/Training: This task should be trained by practice in mock combat situation with consequences for mistakes.

A.17.4 Overturned vehicles

Sensory: Visual sense used primarily.

Perceptual: Must be able to see the position.

Cognitive: This task requires quick reactions as opposed to thinking.

Motor: Turning body vigilantly, holding and firing weapon.

Instructional/Training: This task should be trained by practice in mock combat situation with consequences for mistakes.

A.17.5 Roof tops

Sensory: Visual sense used primarily.

Perceptual: Must be able to see the position.

Cognitive: This task requires quick reactions as opposed to thinking.

Motor: Turning body vigilantly, holding and firing weapon, crawling, balancing.

Instructional/Training: This task should be trained by practice in mock combat situation with consequences for mistakes.

A.18 Firing the individual weapon during movement

A.18.1 Fire around corners while laying flat

Sensory: Primarily visual.

Perceptual: The individual should be able to see around the corner, and will feel force of firing.

Cognitive: This is primarily a motor task.

Motor: Holding weapon, lying on ground, braced against kick.

Instructional/Training: This task should be trained by practice with live ammunition.

A.18.2 Fire around walls; not over walls

Sensory: Primarily visual.

Perceptual: The individual should be able to see around the wall and will feel force of firing.

Cognitive: This is primarily a motor task.

Motor: Holding weapon, braced against kick.

Instructional/Training: This task should be trained by practice with live ammunition.

A.18.3 Fire from within buildings while remaining in the shadows

Sensory: Primarily visual.

Perceptual: The individual should be able to see out the window or door and will feel force of firing.

Cognitive: This is primarily a motor task.

Motor: Holding weapon, braced against kick.

Instructional/Training: This task should be trained by practice with live ammunition.

A.18.4 Fire from both shoulders to take full advantage of available cover and concealment to minimize the amount of the body that is exposed to the enemy

Sensory: Primarily visual.

Perceptual: The individual should be able to see around the corner and will feel force of firing.

Cognitive: This is primarily a motor task.

Motor: Holding weapon on either shoulder, braced against kick.

Instructional/Training: This task should be trained by practice with live ammunition.

Appendix B

Analysis of Training Requirements for Entering Buildings

Analysis of Training Requirements for Entering Buildings

B.1 General

B.1.1 Determine the best point of entry into the building

Sensory: Determination of the best point of entry requires positioning oneself to see multiple entry points. *Vision* is the primary and probably sole operational sensory mode.

Perception: Perception of entry points involves some preconceptions about what constitutes as "entry point." Crawl spaces, burnt walls, etc. may be perceived as entry points by personnel whose perceptions are primed.

Cognitive: The trainee must have a grasp of the overall strategic goal in order to quickly eliminate points of entry that are clearly not "best." The approach to the building must be planned with the operational goal in mind. The individual must have knowledge of the possible consequences of entering by various points: these include general consequences, such as the possibility of someone hiding behind a door, as well as specific consequences resulting from particulars of the situation (i.e., ammunition has been fired from a particular window).

Attention must be paid to important details of the operation, and to relevant structural and strategic features of the building to be entered.

From the cognitive perspective, this is essentially a *judgment* task. The individual gathers information regarding the operational goal, the structure of the building, the history of the building (has it been known to be occupied previously? Are the residents friendly or hostile?) and its surrounding area, strategic aspects of entering the building, aspects of each point of entry, and then forms a judgment: Which point of entry is best?

Motor: This is not a motor task as such. It will usually be necessary for the individual to circumambulate the building under cover, which might require crawling, stooping, running, etc.

Instructional/Training: Training of this task should focus on cognitive features. Trainees should learn to assess the situation, including covert viewing of all points of entry, understanding of strategic goals, knowledge of consequences of various types of entry, and knowledge of building structure.

B.1.2 Determine method to be used and move to the selected point of entry

[Comments: Note that this task consists of two distinct operations, one cognitive and one behavioral]

Sensory: Mode primarily visual. Should be able to see the building and its surrounding area.

Perceptual: Individuals should be able to perceive whether movement from point to point puts them in jeopardy: that is, it is necessary to see, not only the intended position, but those positions that would enable an adversary to see it.

Cognitive: The first part of this task is cognitive. Having determined the best point of entry, the individual must make an informed judgement as to the optimal method of entry. This requires knowledge of appropriate skills needed and a reasonable assessment of the individual's own probability of successfully executing those skills.

This task requires that the individual be capable of assessing the probabilities of adversaries occupying various positions relevant to building entry: this includes the position to which the individual wishes to move, points overlooking that position, and positions within the building itself.

Motor: Skill in executing diverse entry techniques will result in increased options and enhanced performance of relevant behaviors. Moving to the point of entry of an occupied building might place the individual in some danger of ambush or attack; consequently a wide range of motor skills could be called upon.

Instructional/Training: Training for this skill should be two-pronged. *Cognitive skills* require training in various methods of entry, and individuals should understand their personal skill levels (probability of success) for each task. *Motor* skills to be used overlap with skills trained for other military operations and can be learned on obstacle courses, etc.

B.1.3 Check all holes, apertures, doorways, windows, etc. for booby traps before entering

Sensory: This task relies mainly on visual inspection of the building, though other senses such as smell and hearing might be attuned to searching for meaningful stimuli. Vision requires light, and thus operations must be conducted during daylight hours, or an illumination device must be used.

Perceptual: The recognition of features that portend booby traps requires extensive familiarity with the kinds of devices that can be used to this end. Knowledge, in other words, will drive perception. The individual must understand the various kinds of triggering and detonation devices and be able to spot these when they are present.

Perceptual: The individual must possess knowledge of the kinds of objects that can be used in booby traps. Note that booby traps are effective to the extent that they are surprising, or improbable. Effective booby traps might be improvised using materials that are handily available and in a way that has not been demonstrated in a training course. Thus the individual must be able to take the adversary's perspective, judging what the other might have judged to be surprising. The individual must make judgments regarding the adversary's motivations for setting traps, including the length of time the adversary may have occupied the building, the length of time they may intend to stay, whether the adversary knows the individual is following, etc.

Instructional/Training: Training needs to focus on two aspects. First, knowledge must be imparted, and second the individual should become motivated to become vigilant in attending to possible trap cues. The trainee should be shown a wide range of

actual booby traps, how to set them, how to defuse them, and what their consequences are. Viewing of graphically wrenching scenes of the effects of booby traps should motivate reamers to watch for them in the field. Non-traumatic booby trap surprises should further be programmed into VE simulations and field exercises.

B.1.4 Use hand grenades

B.1.4.1 Throw a hand grenade through an upper floor window

Sensory: Visual processes are necessary for detection of the target, estimation of its distance, and evaluation of the accuracy of the throw. But the important sensory aspects of this task comprise proprioceptive and kinesthetic sensing of body position and movement during the act of throwing.

Perceptual: Individuals must position themselves so that they can see the window, but someone inside the window cannot see them without exposing themselves.

Cognitive: This task requires motor skill. The individual has already judged the situation and decided to throw the grenade. Cognitively, the trainer must choose a position such that if the grenade misses the mark it does not bounce back and explode nearby.

Motor: This task is defined as a motor task. Individuals must learn to adjust their musculature to accommodate the weight of the grenade, and to propel the object with accuracy.

Instructional/Training: The training of the motor task almost certainly requires practice with an actual grenade. The individual should repetitively throw the object at a variety of upper floor windows, from various distances, until a criterion hit rate is met.

B.1.4.2 Throw a hand grenade through a lower level window

Sensory: Identical to B.1.4.1.

Perceptual: Same as B.1.4.1, except that the individual, being at the level of the window, will be able to see the target more clearly.

Cognitive: Same as B.1.4.1, except that rather than selecting a position where the grenade will not bounce back, the individual must select a position where the grenade will not be thrown back, and where the debris will not hit.

Motor: Same as B.1.4.1.

Instructional/Training: Same as B.1.4.1.

B-2 Entering upper floors

B.2.1 Use grappling hook

B.2.1.1 Select suitable grappling hook and rope

Sensory: The individual should feel the strength and texture of the rope.

Perceptual: The individual must perceive qualities of the rope and hook, observe the distance to the window or target.

Cognitive: Knowledge of important qualities in hook and rope are necessary. Individuals must judge whether the rope and hook will hold their weight, fray, or break.

Motor: This is a cognitive task.

Instructional/Training: Trainees should be informed about the various qualities that affect the performance of a grappling hook. They should be allowed to hold and examine a variety of ropes and hooks. They should also view (via training film or example) the problems that can arise when inappropriate apparatus is used.

B.2.1.2 Throw the grappling hook, allowing the rope to play

Sensory: The individual must be able to see the target and the progress of the thrown rope. Kinesthetic and proprioceptive cues guide the throwing.

Perceptual: The individual must watch the progress of the hook toward its target and see that the rope hangs freely.

Cognitive: N/A

Motor: This is a motor task. The individual must have practiced this behavior in advance in order to execute it reliably under stressful conditions.

Instructional/Training: Trainees should have hands-on experience throwing various grappling hooks under various conditions. Training should continue until a criterion is met.

B.2.1.3 Pull on the grappling hook to ensure it has a secure hold before beginning to climb

Sensory: Proprioceptors will enable determination of the amount of resistance offered by the rope.

Perceptual: The individual will perceive the amount of weight that can be applied to the rope and the amount of movement of the hook.

Cognitive: Observation of the hook and rope will allow judgment of the security of the attachment.

Motor: Pulling on rope.

Instructional/Training: Individuals should be allowed to pull on ropes that vary in the security of their attachment to the window or ledge. The experience of failing when a hook comes loose might be educational as well.

B.2.1.4 If the grappling hook is secured to a window, pull it to one corner to improve chances of a good “bite” and to minimize exposure of the climber to lower story windows during the climb

Sensory: Must be able to see where the hook is lodged.

Perceptual: The individual must see the corner of the window and the placement of the hook.

Motor: The person must pull on the rope.

Cognitive: The individual must be able to imagine the placement of the unseen side of the hook and judge the security of the bite of the hook in the window or ledge. Further, the individual must imagine climbing the rope and determine whether, from the adversary's point of view, the climbing position will be vulnerable.

B.2.1.5 When the climber is exposed to enemy fire, use smoke and diversionary measures to conceal the climber

- a. When using smoke, consider wind direction and ensure that sufficient amounts of smoke are available.
- b. Diversionary measures that can be used include weapon firing, shouting, and false portrayal of movement.

Sensory: Sees, hears, and smells enemy fire. If smoke is being used as concealment, the individual may be unable to see.

Perceptual: The individual perceives enemy fire, evidence of wind direction, the presence of smoke or other concealment agents.

Cognitive: This task is strategic in nature. The individual must assess the situation from the perspective of the adversary and present a display of diversionary misinformation. If smoke is used to screen, the individual must judge whether it is sufficient in quantity and quality, how long it can be expected to remain, etc.

Motor: It may be necessary to hold one's breath if in smoke, and to release the rope to free a hand for weapon firing. In general, the whole body will be engaged in movement.

Instructional/Training: As the goal of this task is deceptive self-presentation, it would be beneficial for the trainee to watch others attempt to conceal themselves. War games and other field exercises are best for motivating and demonstrating to the trainee the success or failure of various diversionary techniques.

B.2.1.6 Provide cover by friendly fire for movement of marines moving from building to building and for those climbing buildings

Sensory: The individual must not necessarily see the marines who are moving, but must see positions that jeopardize those individuals. Listens for gunfire and indications of concealed enemies.

Perceptual: Must be attentive to position of the marines as well as potential hiding places of adversaries.

Cognitive: Think like the enemy: the individual should distract the attention of adversaries and must be aware that they suspect a cover-up. The individual must be able to imagine the view of the enemy, whether they can see the marines or not.

Motor: Firing weapons, maneuvering for a good view.

Instructional/Training: This behavior should be trained in a variety of realistic adversarial situations.

B.2.1.7 Climber avoids exposure to enemy fires from a lower window

- a. Avoid movement in front of lower windows.

Sensory: Must see windows.

Perceptual: The individual must recognize dangerous lower windows.

Cognitive: Attention to features of the building.

Motor: The individual who is ascending past a window must be able to maneuver to the sides, press body against building out of view of occupants, and contort in whatever way is necessary to prevent being seen by persons within the building.

Instructional/Training: The individual should practice climbing past a variety of windows, with mock adversaries inside trying to fire on him. It does seem likely that a virtual rope would give practice in climbing, but a real rope combined with VE window and enemy fire should result in a realistic training scenario.

- b. Clear lower rooms with hand grenades before ascending outside the window. Loosen the safety pin on the hand grenade to permit one-handed use.

Sensory: Tactile sense indicates when the safety pin has been loosened.

Perceptual: The individual must see the window and whether it is obstructed with a screen, etc. The trainee must loosen the pin, relying mostly on tactile feedback to determine when it is properly set.

Cognitive: The individual must judge whether throwing a grenade through a window is the most efficacious procedure, and must estimate in advance where exploding fragments will go, and how to avoid being hit by them.

Motor: Hanging on a rope loosening a grenade pin by sense of feel, throwing grenade accurately.

Instructional/Training: Because the weight of the grenade and the individual's body on the rope cannot be simulated by software, it is recommended that trainees gain experience loosening real pins from actual grenades while climbing real ropes.

B.2.1.8 Enter the building using the lowest possible silhouette

- a. Hook one leg over the window sill and enter sideways, straddling the ledge.

Sensory: Tactile and kinesthetic senses convey the degree of stability of the window ledge; the individual must look and feel for sharp objects (shards of glass, etc.) and obstacles to entry.

Perception: The individual perceives activity inside the window and aspects of the window sill that might facilitate or hinder entry by straddling.

Cognitive: The individual must judge the best way to enter the building, based on perceptual features of the building face and window, and assessment of the probability of attack from within and without.

Motor: This is a motor task.

Instructional/Training: This task should be trained by climbing over real window sills, which can possibly be embedded in a VE simulation. That is, a sawhorse or other object could substitute for the window sill, providing realistic force feedback, while environmental details are simulated with computer graphics.

- b. Enter head first.

Sensory: Looks into window; listens for cues to presence of threat.

Perceptual: The individual observes characteristics of window and interior, feels sturdiness of sill and sees and feels objects on the floor within.

Cognitive: The individual must assess the best way to enter the building and estimate the risks involved with a head first dive as opposed to another method of entry.

Motor: The individual must perform a complex series of gestures, which will vary with the situation. Some adjustments will be made in mid-motion, as the individual is able to see more of the room during entry.

Instructional/Training: Trainees should dive through windows of various sizes, from various angles.

B.2.2 Use the basic seat-hip rappelling technique to descend from the top of the building into a window

Sensory: Must see the window, feel the wall while rappelling, and sense position of body.

Perceptual: It may be necessary for the individual to perceive height above the ground accurately, or at least not to overestimate it, resulting in panic.

Cognitive: Must assess risks of being exposed to persons within the building and in the surrounding terrain.

Motor: This task requires a complex and unique set of behaviors involved in balancing on the rope and maneuvering toward the target.

Instructional/Training: Rappelling is probably best learned by doing.

B.3 Entering middle floors

B.3.1 Use two-person lift, supported

B.3.1.1 Two marines stand facing each other holding a support (such as a board, bar, weapon) between them

Sensory: This task requires proprioceptive sensing of body position, visual location of the other person and the support.

Perceptual: Must perceive their own and the other's positions relative to the support and the environment.

Cognitive: The marines must decide to execute this action, select an appropriate support, and select which of their group shall go in.

Motor: Individuals must first pick up the support, then hold it with a firm grip, positioned so that the sudden weight of the third person will not upset them.

Instructional/Training: The two-person lift (supported) should be practiced with actual weight of a person.

B.3.1.2 A third marine stands on the support and is lifted upward to the entrance

Sensory: The third marine must be able to balance, sense proprioceptive and kinesthetic cues that indicate failing.

Perceptual: The third person must be able to see the targeted window and notice any changes in the position of the support.

Cognitive: N/A

Motor: The two marines holding the support must remain balanced and exert some strength to hold it steady. The third marine must be able to climb up (dexterity and strength required) and balance on the support.

Instructional/Training: This movement should be practiced by teams of individuals taking turns to get the feeling of holding the support as well as climbing up on it.

B.3.2 Use two-person life, unsupported

B.3.2.1 Two marines stand facing each other with their hands cupped

Sensory: Visual contact is required between the two; tactile sense of the strength of grip.

Perceptual: They must be able to see their surroundings, including the window they have targeted and the individual they are going to boost up to it.

Cognitive: Decide if this is the best method.

Motor: This is a motor task.

Instructional/Training: Should be practiced in teams.

B.3.2.2 A third marine stands on their cupped hands and is lifted upward to the entrance

Sensory: Third marine must see the entrance, and proprioceptively/kinesthetically sense position relative to the entrance and his partners.

Perceptual: They must be able to see their surroundings, including the window they have targeted and the individual they are going to boost up to it.

Cognitive: Judge whether the three individuals are capable of the task, and whether it is the best way to accomplish the goal.

Motor: This is a motor task.

Instructional/Training: This task should be practiced by teams of individuals rotating roles.

B.3.3 Use two-person lift, heels raised

Sensory: Third marine must see the entrance, and proprioceptively/kinesthetically sense position relative to the entrance and his partners.

Perceptual: They must be able to see their surroundings, including the window they have targeted and the individual they are going to boost up to it.

Cognitive: Judge whether the three individuals are capable of the task, and whether it is the best way to accomplish the goal.

Motor: This is a motor task.

Instructional/Training: This task should be practiced by teams of individuals rotating roles.

B.3.4 Use one-person lift

Sensory: Same as previous.

Perceptual: Same as previous.

Cognitive: Same as previous.

Motor: This requires more strength from both individuals than do the previous methods. One individual must bend and support the other's weight, while the second individual must be able to climb up.

Instructional/Training: Individuals should learn this task by practicing it in teams.

B.3.5 Use two-person pull

Sensory: Task requires visual access to the individuals outside the window, as well as proprioceptive/kinesthetic sensation and balance.

Perception: The individual doing the pulling must perceive the other(s), must perceive the location of his/her center of gravity, and must feel the mutual grip strength.

Cognitive: Same as previous.

Motor: This requires strength as well as sense of balance.

Instructional/Training: This task should be trained by practicing it in teams.

B.3.6 Use sling lift

Sensory: Same as previous.

Perceptual: Same as previous, except the individual holding the weapon should see that it is not pointed at anyone.

Cognitive: Same as previous.

Motor: Same as previous,

Instructional/Training: Same as previous.

B.4 Entering ground level floors

B.4.1 If doors, windows, or existing holes are used, check for and clear booby traps

Sensory: This task relies mainly on visual inspection, though other senses such as smell and hearing might be attuned to searching for meaningful stimuli. Vision requires light, and thus operations must be conducted during daylight hours, or an illumination device must be used.

Perceptual: The recognition of features that portend bobby traps requires familiarity with the kinds of devices that can be used to this end. Knowledge, in other words, will drive perception. The individual must understand the various kinds of triggering and detonation devices and be able to spot these when they are present.

Cognitive: The individual must possess knowledge of the kinds of objects that can be used in booby traps. Note that booby traps are effective to the extent that they are surprising, or improbable. Effective booby traps might be improvised using materials that are handily available, and in a way that has not been demonstrated in a training course. Thus the individual must be able to take the adversary's perspective, judging what the other might have judged to be surprising. The individual must make judgments regarding the adversaries motivations for setting traps, including the length of time the adversary may have occupied the building, the length of time they may intend to stay, whether the adversary knows the individual is following, etc.

Instructional/Training: Training needs to focus on two aspects. First, knowledge must be imparted, and second the individual should become motivated to become vigilant in attending to possible trap cues. The trainee should be shown a wide range of actual booby traps, how to set them, how to defuse them, and what their consequences are. Viewing of graphically wrenching scenes of the effects of booby traps should motivate reamers to watch for them in the field. Non-traumatic booby trap surprises should further be programmed into VE simulations and field exercises.

B.4.2 Use demolition, smaw's, tanks, artillery, or other weapons to blast an entrance way

B.4.2.1 Tie or tape 9 blocks of C-4 on a pole and place it against the target between waist and chest height and detonate. take protective measures to keep blast from injuring friendly marines

Sensory: The individual must be able to see and feel the equipment and objects being used.

Perceptual: The trainee must see the location against which the C-4 is placed. The individual must also not perceive the explosion itself, unless from a great distance away.

Cognitive: Trainees must be able to estimate the approximate result of the explosion of 9 blocks of C-4 in order to determine where to place the substance for best effect, as well as where to position themselves to avoid injury.

Motor: The individual must tape the C-4 to the pole, must place the pole, and should run 300 meters, or to another building, to escape the explosion.

Instructional/Training: Trainees should be shown movies of a number of explosions under various situations, in order to familiarize them with the methods involved before they are allowed to try blasting an entrance way themselves. Because of the danger of the explosive, prudent cautionary techniques should be taught.

B.4.2.2 Metal reinforcing bars will not be cut with C-4 once exposed, use tnt for cutting metal reinforcing bars

Sensory: Must see the hole in the wall clearly enough to determine where to place the TNT. The individual must hear the explosion in order to know when to return.

Perceptual: The individual must perceive the hole well enough to recognize the presence of reinforcing rods and to know where the TNT should be placed in order to remove them. As previously stated, the individual will not perceive the explosion itself, unless from distance.

Cognitive: The individual must estimate the thickness of the bars and apply one pound of TNT if the diameter is one inch or less, two pounds if it is between 1 and 2 inches in diameter.

Motor: The individual must Place the TNT and detonating device and run.

Instructional/Training: Same as previous.

B.4.3 Entering through doors

B.4.3.1 Check for booby traps

Sensory: This task relies mainly on visual inspection of the building, though other senses such as smell and hearing might be attuned to searching for meaningful stimuli. Vision requires light, and thus operations must be conducted during daylight hours or an illumination device must be used.

Perceptual: The recognition of features that portend booby traps requires extensive familiarity with the kinds of devices that can be used to this end. Knowledge, in other words, will drive perception. The individual must understand the various kinds of triggering and detonation devices and be able to spot these when they are present.

Cognitive: The individual must possess knowledge of the kinds of objects that can be used in booby traps. Note that booby traps are effective to the extent that they are surprising or improbable. Effective booby traps might be improvised using materials that are handily available, and in a way which has not been demonstrated in a training course. Thus the individual must be able to take the adversary's perspective, judging what the other might have judged to be surprising. The individual must make judgments regarding the adversaries motivations for setting traps, including the length of time the adversary may have occupied the building, the length of time they may intend to stay, whether the adversary knows the individual is following, etc.

ilant in attending to possible trap cues. The trainee should be shown a wide range of actual booby traps, how to set them, how to defuse them, and what their consequences are. Viewing of graphically wrenching scenes of the effects of booby traps should motivate reamers to watch for them in the field. Non-traumatic booby trap surprises should further be programmed into VE simulations and field exercises.

B.4.3.2 Line up on side of least resistance

- a. For doors that swing inward, line up on the hinge side.
- b. For doors that open outward, line up on the doorknob side.

Sensory: Individuals must be able to see the door, the area outside the door, and the positions of their team mates.

Perceptual: Must recognize the hinge and doorknob and determine whether the door swings in or out.

Cognitive: Must decide on the order for lining up, and on which side. Individuals must be able to imagine a large range of possible consequences of opening the door, including probability estimates for each consequence, risks associated with each, and costs of risks.

Motor: Walk or run to line up beside the door.

Instructional/Training: The determination of which side to line up on can be communicated verbally and through graphics and movies.

B.4.3.3 Position members of the assault team as close as possible to the door

Sensory: Individuals must be able to see the door, the area outside the door, and the positions of their team mates.

Perceptual: Must recognize the hinge and doorknob, and determine whether the door swings in or out.

Cognitive: Must decide on the order for lining up, and on which side. Individuals must be able to imagine a large range of possible consequences of opening the door, including probability estimates for each consequence, risks associated with each, and costs of risks.

Motor: Walk or run to line up beside the door.

Instructional/Training: Practice in the field should train this behavior. The reasons for this method should be explained to trainees, and possible consequences of improper technique should be described, shown graphically, and perhaps experienced by them.

B.4.3.4 Hold weapons in the firing hand.

Sensory: Must feel the weight of the weapon and see it.

Perceptual: Must see where the weapon is aimed, who might be in the line of fire, and its position relative to the door to be entered.

Cognitive: Trainees should have knowledge of their weapons, and should be attentive to the door and the positions of comrades in case it is necessary to aim and fire.

Motor: This is a motor task.

Instructional/Training: The holding and positioning of weapons appropriately for various situations should be trained by repetition from boot camp onward.

B.4.3.5 When using the stealth approach, team members use the “thumb-back, squeeze-up” for hand signals.

Sensory: Team members must be able to see one another.

Perceptual: Individuals must be able to recognize hand signals given by the others.

Cognitive: This task requires vigilance and knowledge of the operation of the hand signals.

Motor: Trainees move their fingers to indicate movements.

Instructional/Training: Individuals should be familiar with the hand signals, including being able to make them and interpret them.

Appendix C

Analysis of Training Requirements for Clearing Rooms

Analysis of Training Requirements for Clearing Rooms

C.1 General

C.1.1 Gain surprise

Sensory: All senses may be involved.

Perceptual: As the trainees cannot see into the room, they must listen for any sounds and examine any signs outside the room that might indicate the presence of persons inside.

Cognitive: While the trainees are inducing surprise, they also will experience surprise, as they will not know what to expect. Thus attention must be focused, memory or any information regarding the interior and probable occupation of the room must be available for recall, and perceptions must be primed for any indication of danger or counter-attack.

Motor: Probably requires quick movements, such as opening a door rapidly and firing into the room, or kicking door open, etc.

Instructional/Training: To train this task, individuals should experience the uncertainty of surprising a room without knowing what is inside. Mock combat should be employed so that consequences are delivered for inept performance of the task.

C.1.2 Clear room quickly

Sensory: Primarily visual and auditory.

Perceptual: Must see if anyone is in the room, including hidden corners. etc.

Cognitive: This task requires some knowledge of hiding places.

Motor: Running, walking, possibly firing weapon.

Instructional/Training: This should be practiced in exercises and mock combat.

C.1.3 Employ hand grenades, rifle fire, and speed to dominate the room, overwhelm and keep the enemy off guard

Sensory: All senses are involved.

Perceptual: Must be able to see and/or hear targets of grenade and rifle fire, and be able to select a position where the individual will not be injured by explosions.

Cognitive: Must judge a best direction to throw grenades or fire, which includes recall of any information about the contents and likely occupation of the room.

Motor: Throwing and jumping back, and firing a weapon, while ducking, running, and evading counter fire.

Instructional/Training: This task should be trained in mock combat with consequences.

C.1.4 Vary techniques to make it more difficult for the enemy to prepare a defense

Sensory: All senses are potentially involved.

Perceptual: Must be able to see and/or hear presence of enemy and positions that allow varied techniques.

Cognitive: This task describes a strategy for surprising the enemy by producing unexpected behavior. Knowledge of a range of techniques is necessary, as well as an understanding of what techniques would be expected.

Motor: Might involve throwing, running, kicking, shoving, etc.

Instructional/Training: This task should be learned in interaction with another, with the goal of surprising the other.

C.2 Assign sectors of fire

Sensory: Visual and auditory.

Perceptual: The individual must be able to see the area and to see what team members are present and available to fire. Others must be able to hear verbal commands.

Cognitive: The individual must judge the relative abilities of team members, based on personal information and position, to apply fire to particular sectors.

Motor: Vocalize, motion with arms.

Instructional/Training: This management skill is learned by practice leading a team, especially in mock combat situations.

C.3 Eliminate the threat

C.3.1 If only the enemy is located in the room, may use grenades

Sensory: Visual and auditory senses are involved.

Perceptual: The individual must be able to see the entire room in order to ascertain that only enemies are present.

Cognitive: Knowledge of who is in the room is required, as well as judgment of the appropriateness of using a grenade.

Motor: Look, throw grenade, and stand back.

Instructional/Training: Scanning the room is trained separately from tossing grenades.

C.3.2 If non-combatants are present look for:

1. Weapons in their hands.
2. Threatening actions such as drawing or reaching for a weapon.
3. Positive identification such as uniform or other types of dress.

Sensory: Visual sense is involved.

Perceptual: Must be able to see dress and actions of individuals in the room.

Cognitive: This task requires quick judgment under risk, based on incomplete information. It includes use of knowledge of local customs of dress, behavior, and opinion in order to assess the probability of hostility. It also requires knowledge of possible consequences of error, including both errors that allow surprise attack by non-uniformed individuals and erroneous execution of nonmalicious locals.

Motor: This is a perceptual/cognitive task.

Instructional/Training: Individuals should be trained in interaction with locals and should be trained to identify indications of concealed weapons.

C.4 Control the situation and personnel

C.4.1 Speak to live people in the room with a loud commanding voice, using short, to-the-point language

Sensory: Kinesthetic and auditory senses.

Perceptual: Must listen to one's own voice.

Cognitive: The individual must understand the status relationship between self and others, as it is perceived by others, and how to control the interaction so as to induce compliance.

Motor: Loud speaking

Instructional/Training: This could be trained by having the individual order a team around, or by practicing with a coach present.

C.4.2 Search the dead. Perform an "eye thump" on the body that will cause a deep pain response on a live person

Sensory: Kinesthetic, auditory, and visual senses are engaged.

Perceptual: The individual must see the body well enough to detect movement such as breathing, and must hear if the allegedly deceased person groans.

Cognitive: The judgment of whether the person is dead or alive requires some knowledge of vital signs; the individual must also know how to perform the “eye thump,” and distinguish reaction from momentum.

Motor: Bending, walking, and performing of “eye thump” are included in this task.

Instructional/Training: The “eye thump” should be demonstrated by an instructor and practiced on a dummy. Search methods should likewise be demonstrated and practiced.

C.4.3 Search the room:

1. Identify other threats.
2. Identify booby traps.
3. Search live personnel at this time.

Sensory: This task is primarily visual.

Perceptual: The individual must be able to see objects in the room, including cues to the presence of booby traps, etc.

Cognitive: The individual requires knowledge of booby trap techniques and other potential threats.

Instructional/Training: Trainees should be instructed in methods for triggering booby traps and in techniques for rapidly searching a room, including live personnel. They should also practice rapidly searching a room with consequences for missing something.

C.5 If the purpose of the mission was to secure and evacuate personnel or equipment, do that at this time while maintaining rear security

Sensory: All senses are involved.

Perceptual: Individuals must be able to see and move personnel or equipment that are to be secured and evacuated,

Cognitive: The goals of the mission must be known and methods for securing and evacuating personnel and equipment must be applied.

Motor: Could be any motor movements.

Instructional/Training: This should be taught as a strategy.

C.6 Organization

C.6.1 Covering party places fire on critical areas to support the search party making the entrance

Sensory: Primarily visual and auditory.

Perceptual: The individual must see the position of the search party as well as critical areas to fire upon.

Cognitive: Must judge which areas are critical.

Motor: Walking, running, firing weapons.

Instructional/Training: This task should be learned in mock combat.

C.6.2 Search team enters and clears the building or room while the covering party provides protective fires

Sensory: Visual/auditory.

Perceptual: Individuals clearing the room must see what they are doing and hear the approach or presence of enemies. Those providing protection must be able to see the area.

Cognitive: This requires judgments of where enemies might be located and knowledge of techniques for clearing buildings and providing cover.

Motor: Walking, running, firing weapons.

Instructional/Training: This task should be learned in mock combat.

C.7 Clearing the room

C.7.1 Throw a cooked-off grenade into the room. (Pull the pin, cock the arm into the throwing position, release the safety lever, count "1000-1, 1000-2," then throw the grenade into the room.)

Sensory: Visual, auditory, tactile, and kinesthetic senses are involved.

Perceptual: The individual feels the grenade in the hand, sees the pin, and tenses muscles to throw.

Cognitive: The person must estimate the passage of time and be quite certain of where the grenade will end up exploding.

Motor: Pull pin with fingers, extend arm to throwing position, hold pose while counting, then throw.

Instructional/Training: This could be trained with a dummy grenade, or with a weighty prop such as a rock, in a VE with graphic explosions etc.

C.7.2 Enter the room immediately after detonation; spray the room with three round bursts of automatic weapons fire as the room is entered

Sensory: Kinesthetic, visual, auditory senses involved.

Perceptual: The person must be able to see the room while firing into it.

Cognitive: The element of surprise is called for. The individual is not certain what will be found upon entering the room and must be very attentive to the possibility of danger.

Motor: Running, jumping, ducking, firing weapons.

Instructional/Training: Mock combat, real or virtual.

C.7.3 Assume a position where the entire room can be seen and, if the room is clear, call out "clear." If the room is not clear, eliminate the threat and then announce "clear."

Sensory: Visual, auditory, etc.

Perceptual: Must be able to see the room and any threats in it.

Cognitive: Must judge whether threats are present.

Motor: Running, jumping, ducking, crouching, shouting.

Instructional/Training: Should be trained in mock combat, real or virtual.

C.7.4 Other marines call out "coming in, last name" and then enter the room to take up firing positions and search it for concealed enemy personnel. When the search is completed, call out "clear"

Sensory: Primarily auditory.

Perceptual: Individuals must be able to see one another and the objects within the room, and to hear one another.

Cognitive: Must pay attention to information delivered by those entering the room and strategically and thoroughly search the room.

Motor: Running, jumping, ducking, crouching, shouting.

Instructional/Training: Should be trained in mock combat, real or virtual.

C.7.5 As other marines enter the room, they call out “coming in, last name” as they enter the room

Sensory: Primarily auditory.

Perceptual: Individuals must be able to see one another and the objects within the room, and to hear one another.

Cognitive: Must pay attention to information delivered by those entering the room.

Motor: Running, jumping, ducking, crouching, shouting.

Instructional/Training: Should be trained in mock combat, real or virtual.

C.7.6 Marines are directed to a safe position from where they can conduct a safe search of the entire building

Sensory: Primarily visual.

Perceptual: Individuals must be able to see the safe position and visually search the building.

Cognitive: Someone must be giving orders, and some others must receive them. Thus, one person is making decisions with judgments of risk, while others are complying.

Motor: Walking, running, crouching, ducking, etc.

Instructional/Training: This should be well practiced in mock combat.

C.7.7 Leave doors open after clearing rooms and make a predetermined mark on the door frame to signify that the room has been cleared

Sensory: Primarily visual.

Perceptual: Must be able to see marks on doors.

Cognitive: Must have knowledge of what the mark is.

Motor: Open doors, make mark.

Instructional/Training: This task should be explained, and perhaps practiced on a real door.

C.8. Use C-4, claymore mines, and TNT demolitions to gain access to rooms

Sensory: Visual, auditory.

Perceptual: Individuals must be able to see what they are doing and will hear explosions.

Cognitive: The individual must have knowledge of explosive techniques and must also be able to judge whether this method is preferable over less destructive ones.

Motor: Using explosives, running, ducking, etc.

Instructional/Training: Individuals should receive classroom instruction in the use of explosives, augmented with experience with actual explosives.

C.9 Clear rooms with closed doors

C.9.1 Check for booby traps, do not use door knobs

Sensory: Visual.

Perceptual: Looks for evidence of booby traps.

Cognitive: Possesses knowledge of booby traps, and recalls this.

Motor: Varies.

Instructional/Training: Should be trained in booby trap methods and practiced at opening doors without using the knobs.

C.9.2 Use demolitions or kick open door

Sensory: Visual, auditory, kinesthetic.

Perceptual: The individual must see the door and observe its response to force.

Cognitive: This is primarily a motor task.

Motor: Apply explosives to door or kick foot.

Instructional/Training: Individuals should practice this behavior on a real door.

C.10.2 Check for sewers and tunnels

Sensory: Visual.

Perceptual: The individual must be able to see all parts of the room.

Cognitive: The individual must have knowledge of cues to the presence of sewers and tunnels.

Motor: Walking, looking.

Instructional/Training: Classroom instruction should teach individuals where to look for possible sewers and tunnels.

- C.11** **Avoid hallways if possible; if movement along hallways is necessary, move along the side of walls as quickly as possible to get out of the hallway**

Sensory: Primarily visual.

Perceptual: Individuals must be able to see where they are.

Cognitive: Must be able to judge alternate routes.

Motor: Walking, running, moving quickly along side of walls.

Instructional/Training: The rule should be imparted in classroom instruction/text and should be practiced in mock combat.

- C.12** **Move between floors**

- C.12.1** **Check stairways for booby traps and enemy fire**

Sensory: This task is primarily visual.

Perceptual: The individual must be able to see objects in the stairways, including cues to the presence of booby traps, enemy fire, etc.

Cognitive: The individual requires knowledge of booby trap techniques and other potential threats.

Instructional/Training: Trainees should be instructed in methods for triggering booby traps.

- C.12.2** **If necessary, use explosives or heavy weapons to move between floors by blowing holes in ceiling, floors, or walls**

Sensory: Primarily visual, auditory.

Perceptual: Must see to place explosives or fire weapons.

Cognitive: Requires knowledge of explosives and weapons, and judgment of possible routes through ceiling, floors, or walls.

Motor: Firing heavy weapons or placing explosives, running, ducking.

- C. 13** **Mark the building once it has been completely secured and announce "all secure"**

Sensory: Primarily visual.

Perceptual: Must have seen the entire interior of the building and see where to make mark.

Cognitive: The individual must be competent to make a judgment that the building is secure.

Motor: Marking, shouting.

Instructional/Training: This skill should be trained in a classroom and practiced in mock combat.

C.14. Reorganize the force:

1. Replenish and/or redistribute ammunition
2. Treat wounded
3. Evacuate wounded

Sensory: Visual, auditory, etc.

Perceptual: Must see and communicate with team members.

Cognitive: Individuals must understand the reorganization plan and agree to it.

Motor: Walking, carrying objects and personnel, transferring burdens between individuals.

Instructional/Training: This should be learned in classroom and practiced in mock combat and exercises.

Appendix D

Analysis of Training Requirements for Establishing Defensive Positions

Analysis of Training Requirements for Establishing Defensive Positions

D.1 Fundamental skills

D.1.1 Preparation of buildings

1. Selection of weapon positions.
2. Preparation of weapon positions such as window positions and loopholes.
3. Secure and fortify doors, hallways, stairs, windows, floors, ceilings, unoccupied rooms, basements, upper floors and roofs.
4. Select and prepare interior routes in building.
5. Take fire prevention measures.
6. Establish communications within the building and with other forces outside the building.
7. Establish obstacles.
8. Establish fields of fire.

Sensory: All senses are involved.

Perceptual: This task comprises many perceptions.

Cognitive: This set of fundamental skills requires judgments of risk and quite a lot of knowledge which can be easily recalled.

Motor: Many movements are used.

Instructional/Trainingg: These skills should be taught in classroom and readings, and then practiced in exercises and mock combat.

D.1.2 Select and prepare tank/APC positions

Sensory: Primarily visual.

Perceptual: The individual must be able to see the terrain.

Coenitive: Individuals must have knowledge of what comprises a good tank or APC position including strategic offensive and defensive considerations.

Motor: Driving vehicle, running and walking, possibly lifting objects to clear a way and to camouflage/conceal the vehicle.

Instructional/Training: Selection and preparation aspects should be taught in text and classroom, and the task should be practiced in exercises and mock combat.

D.1.3 Select and prepare ATGM positions

Sensory: Primarily visual.

Perceptual: The individual must observe enemy tank positions and/or likely routes to be used soon.

Cognitive: Knowledge of operating the ATGM as well as of what constitutes a vulnerable tank position is required.

Motor: Walking, running, crouching, carrying heavy equipment.

Instructional/Training: Selection and preparation aspects should be taught in text and classroom, and the task should be practiced in exercises and mock combat.

D.1.4 Select and prepare sniper positions

Sensory: Primarily visual.

Perceptual: The individual must observe actual and/or potential enemy positions.

Cognitive: Snipers should be able to identify optimal strategic positions; they should also acquire skill with weapons.

Motor: Walking, running, crouching, carrying heavy equipment.

Instructional/Training: Selection and preparation aspects should be taught in text and classroom, and the task should be practiced in exercises and mock combat.

D.2 Identify and take up hasty firing positions such as:

D.2.1 Corners of buildings

Sensory: Primarily visual.

Perceptual: Must see building corners and safe routes to them.

Cognitive: The individual should retain a mental model or map of the area, as well as be able to judge good positions.

Motor: Running, crouching, crawling, possibly carrying heavy equipment.

Instructional/Training: This skill should be practiced in exercises and mock combat.

D.2.2 Behind a wall

Sensory: Primarily visual.

Perceptual: Must see walls and safe routes along them.

Coenitive: The individual should retain a mental model or map of the area. as well as be able to judge good positions.

Motor: Running, crouching, crawling, possibly carrying heavy equipment.

Instructional/Training: This skill should be practiced in exercises and mock combat.

D.2.3 Windows and doorways

Sensory: Primarily visual.

Perceptual: Must see windows and doonvays and safe routes to them.

Cognitive: The individual should retain a mental model or map of the area, as well as be able to judge good positions.

Motor: Running, crouching, crawling, possibly carrying heavy equipment.

Instructional/Training: This skill should be practiced in exercises and mock combat.

D.2.4 Unprepared loopholes

Sensory: Primarily visual.

Perceptual: Must see unprepared loopholes and safe routes to them.

Coanitive: The individual should retain a mental model or map of the area, as well as be able to judge good positions.

Motor: Running, crouching, crawling, possibly carrying heavy equipment-

Instructional/Training: This skill should be practiced in exercises and mock combat.

D.3 Prepare a defensive fighting positions

D.3.1 Barricaded windows

Sensory: Primarily visual.

Perceptual: Barricaded window must allow the individual to see through it to the street.

Coenitive: The individual should have a conception of how vulnerable/visible the window is to enemies.

Motor: Running, walking, lifting, holding and possibly firing weapon.

Instructional/Training: This skill should be taught in classroom and text, and practiced in exercises and mock combat.

D.3.2 Fortified loopholes

Sensory: Primarily visual.

Perceptual: Loophole must allow the individual to see the area.

Cognitive: The individual should have a conception of how vulnerable/visible the fortified loophole is to enemies.

Motor: Running, walking, lifting, holding and possibly firing weapon.

Instructional/Training: This skill should be taught in classroom and text, and practiced in exercises and mock combat.

D.3.3 Dig and fortify holes on first floor, preferably next to a wall

Sensory: Primarily visual.

Perceptual: Hole must allow individuals to see the area more easily than they can be seen by enemies.

Cognitive: The individual should have a conception of how vulnerable/visible the hole is to enemies.

Motor: Running, walking, digging, lifting, holding and possibly firing weapon.

Instructional/Training: This skill should be taught in classroom and text, and practiced in exercises and mock combat.

D.3.4 Knock holes in walls just large enough for observation and fields of fire

Sensory: Primarily visual.

Perceptual: Holes must allow individuals to see the area more easily than they can be seen by enemies.

Cognitive: The individual should have a conception of how vulnerable/visible the holes are to enemies.

Motor: Running, walking, pounding, holding and possibly firing weapon.

Instructional/Training: This skill should be taught in classroom and text, and practiced in exercises and mock combat.

D.4 Establish communications with other members of the defensive team via wire or radio

Sensory: Primarily auditory.

Perceptual: Must be able to hear wire or radio communications.

Cognitive: Knowledge of operation of instruments is required, as well as understanding of the likelihood transmissions are being intercepted.

Motor: Holding equipment, tuning knobs.

Instructional/Training: Individuals should learn to operate communications equipment in classrooms, by text, and with practice.

D.5 Establish defensive fires for different weapons for assigned sectors

D.5.1 Machine-gun positions

1. Place on lower floors to provide grazing fires to cover streets, open areas, and avenues of approaches.
2. Keep weapons within the building and in shadows.

Sensory: Primarily visual.

Perceptual: Must be able to see large areas of the environment.

Cognitive: Must be able to imagine the visibility of positions from the enemy's perspective.

Motor: Hold and fire weapons, standing.

Instructional/Training: This skill should be trained in mock combat and exercises.

D.5.2 Anti-Armor Weapons

1. Position these weapons on upper floors to give them longer ranges of fire, and to enable them to fire at the tops and flanks of personnel carriers and tanks.
2. Maintain a clear view of at least:
 - 12 by 15 feet behind TOW, DRAGON, and SMAW weapons
 - 4 feet behind the LAW.
3. Rooms containing these weapons must have open areas of ventilation of at least 21 square feet (a 3 by 7 foot doorway meets this requirement).

4. All personnel in the room must wear ear plugs and cannot be positioned in the rear of the room.

Sensory: Primarily visual.

Perceptual: Individuals must be able to see the surrounding area as well as the room they are set up in.

Cognitive: Individuals must understand the operation of these weapons, including rationale for ventilation and other special requirements.

Motor: Walking, running, crouching, firing large weapons.

Instructional/Training: Use of these weapons should be trained in classrooms and with text, as well as practiced in exercises

D.5.3 Fortify all positions to include overhead cover for those located on roof tops

Sentry: Visual and kinesthetic senses are involved.

Perceptual: Must be able to see materials to use for fortification, as well as to see surrounding area, after fortifications are in place.

Cognitive: Individuals must have knowledge of fortification techniques.

Motor: Walking, crouching, lifting, pounding, etc.

Instructional/Training: These skills should be taught in classroom/text, and then practiced in exercises and mock combat.

D.6 Employ integrated obstacles, barriers, and defensive fires

D.6.1 Emplace mines and booby traps

Sensory: Primarily visual.

Perceptual: Must be able to see materials and positions.

Cognitive: Requires retrievable knowledge of techniques for placing mines and making booby traps.

Motor: Walking, crawling, crouching, running, lifting, digging.

Instructional/Training: Information should be imparted in classrooms and text; these skills should also be practiced in exercises and mock combat.

D.6.2 Overturn vehicles in street area approaching the building

Sensory: Primarily visual.

Perceptual: Must be able to see vehicles and surrounding area.

Cognitive: Must estimate vulnerability/visibility to enemies,

Motor: Walking, crawling, crouching, running, lifting, pushing.

Instructional/Training: These skills should be practiced in exercises and mock combat.

D.6.3 Pile rubble and debris along avenues of approaches

Sensory: Primarily visual.

Perceptual: Must be able to see debris and surrounding area.

Cognitive: Must estimate vulnerability/visibility to enemies.

Motor: Walking, crawling, crouching, running, lifting, pushing.

Instructional/Training: These skills should be practiced in exercises and mock combat.

D.6.4 Dig tank traps

Sensory: Primarily visual.

Perceptual: Must be able to see tank traps and surrounding area.

Cognitive: Must estimate vulnerability/visibility to enemies.

Motor: Walking, crawling, crouching, running, digging, lifting, pushing.

Instructional/Training: These skills should be taught in classroom and text, and practiced in exercises and mock combat.

D.6.5 Cover obstacles and barriers with direct and indirect fires

Sensory: Primarily visual.

Perception: Must be able to see targets.

Cognitive: This is a strategic maneuver; movement of enemies through the area must be anticipated.

Motor: Hold and fire weapons.

Instructional/Training: This skill should be taught in classrooms and text, and practiced in exercises and mock combat.

Appendix E

Analysis of Fundamental Skills that are Common to all MOUT Tasks

Analysis of Fundamental Skills that are Common to all MOUT Tasks

E.1 Throw grenades

Sensory: Primarily proprioceptive.

Perceptual: The individual must be able to see the target and the effects of grenade explosion.

Cognitive: This is primarily a motor task.

Motor: This task requires timing the release of the grenade, propelling it accurately toward the target, and moving to avoid the blast.

Instructional/Training: Because of the importance of proprioceptive sensation, this task should be practiced using a real object of approximately the weight of a grenade.

E.2 Use camouflage techniques

Sensory: Camouflage itself is usually a form of visual deception, though use of camouflage materials requires all sense modalities.

Perceptual: The individual must be able to discriminate aspects of the environment, which can be exploited in the interest of camouflage.

Cognitive: Knowledge of camouflage techniques must be imparted, recalled at the time it is needed, and judgments must be made regarding the adaptation of materials.

Motor: This is primarily a cognitive task.

Instructional/Training: Knowledge of techniques should be imparted pedagogically, and practiced with real materials. This task is not suited to VE practice, where representational appearances are essentially arbitrary.

E.3 Booby traps

1. Detect booby traps
2. Avoid booby traps
3. Report booby traps

Sensory: The individual must be able to see signs of booby traps.

Perceptual: Knowledge of booby trap techniques will guide the individual's perceptual search.

Motor: This is primarily a perceptual task.

Instructional/Training: If individuals learn to make booby traps of various kinds, through instruction and practice with materials, they will also learn where to look for them.

E.4 Mines

1. Detect mines.
2. Use mines.
3. Arm and disarm mines.

Sensory: Primarily visual.

Perceptual: The detection of mines is basically visual.

Cognitive: The individual must have readily available knowledge regarding cues to the presence of mines and how to disarm them.

Motor: Emplacement of mines requires some digging, bending, etc.

Instructional/Training: Experience with real mines should be integrated with practice in a VE mine field.

E.5 Use obstacles

Sensory: Primarily visual and proprioceptive senses.

Perceptual: The individual must see the relations between the obstacles and other features of the environment, such as the positions of enemies, paths to cover, etc.

Motor: This task can involve lifting, pushing, pulling, climbing, digging, etc.

Instructional/Training: VE training with workarounds can simulate this situation adequately for marines to acquire practice.

E.6 Talk with other members of the team involved in MOUT

Sensory: Primarily auditory.

Perceptual: The marine must be able to perceive the others, either through hearing or vision.

Motor: Talking is a form of motor behavior.

Instructional/Training: This should be practiced in real and virtual situations until verbal communication becomes habitual.

Distribution List

Defense Technical Information Center (4)